AFIT/GLM/LSM/92S-24

A METHODOLOGY FOR MODEL COMPARISON USING THE THEATER SIMULATION OF AIRBASE RESOURCES AND ALL MOBILE TACTICAL AIR FORCE MODELS

THESIS

Heston R. Hicks, Captain, USAF

Lawrence L. Long, Civilian, USAF

AFIT/GLM/LSM/92S-24

The views expressed in this thesis are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

AFIT/GLM/LSM/92S-24

A METHODOLOGY FOR MODEL COMPARISON USING THE THEATER SIMULATION OF AIRBASE RESOURCES AND ALL MOBILE TACTICAL AIR FORCE MODELS

THESIS

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Logistics Management

Heston R. Hicks, B.S.

Captain, USAF

Lawrence L. Long B.S.

Civilian, USAF

September 1992

Approved for public release; distribution unlimited

Preface

This study compares the Theater Simulation of Airbase Resources model to the sortie generation module of the All Mobile Tactical Air Force model and concurrently develops a methodology for comparing models. The research expands on earlier efforts, adding a qualitative comparison and a more rigorous quantitative comparison. The results reveal notable qualitative and quantitative differences between the models. Further research is needed to determine the cause of the quantitative differences. The qualitative differences are believed due primarily to differences in the models' designed fidelity. The methodology employed provides a useful framework for subsequent model comparisons and is refined to improve its future usefulness.

We are indebted to our faculty committee, Lt Col Phillip Miller and Lt Col David Diener for their guidance and perseverance and to Mark Speed and James Klosterboer of Ball Systems Engineering Division for their assistance and technical support. This project would not have been completed without the support provided by the AFIT system operators, particularly Joe Hamlin, Anthony Schooler, Doug Burkholder, and Jack Phillips. We extend our sincere appreciation to Mr. Eric Werkowitz, for his guidance on this project and his interest in the results. Finally, albeit most importantly, we extend our sincerest appreciation to our families for their support and understanding as we sacrificed their time and attention pursuing this effort.

Heston R. Hicks

Lawrence L. Long

Table of Contents

Page

List of Figures

List of Tables

AFIT/GLM/LSM/92S-24

Abstract

This research compares the Theater Simulation of Airbase Resources (TSAR) model to the sortie generation (SORGEN) module of the All Mobile Tactical Air Force (AMTAF) model, qualitatively and quantitatively, while concurrently developing and proving a model comparison methodology. The qualitative analysis compares the models' background and documentation, features and databases, and useability. The quantitative analysis statistically compares the models' estimates of sorties generated. For the quantitative study, eight variables are chosen and assigned high and low values for use in a 2^8 1/4 fractional factorial experimental design. Equivalent input databases are developed from a TSAR F-15C database and pilot trials are run to test the factor levels and assess variability. Finally, 64 experimental trials are run and paired differences of the results are tested to determine the statistical equivalence of the models. Results reveal notable differences in the models, both qualitative and quantitative. Further research is needed to analyze the quantitative differences. The qualitative differences are believed due primarily to differences in the models' designed fidelity. The methodology developed provides a functional framework for model comparison and is improved for use in future research.

A METHODOLOGY FOR MODEL COMPARISON USING THE THEATER SIMULATION OF AIRBASE RESOURCES AND ALL MOBILE TACTICAL AIR FORCE MODELS

I. Introduction

Issue

The Theater Simulation of Airbase Resources (TSAR) simulation model is currently used by the Air Force Center for Studies and Analysis (AFCSA) and the Munitions Development Branch of Aeronautical Systems Center (ASC/YQ). The All Mobile Tactical Air Force (AMTAF) simulation model, which the Mission Area Planning Section of Aeronautical Systems Center (ASC/XRS) procured to improve its mission area planning capability, and is purportedly easier to use, sees limited use by a few organizations. Both models possess the ability to simulate the capability of an airbase to generate and sustain sorties under wartime conditions, but the estimates produced by the TSAR model are used and trusted while little is known about the capability of AMTAF to produce similar data. To what extent are the two models similar? How can AMTAF and TSAR be compared to determine the extent to which they are equivalent? These questions are of direct concern to ASC/XRS and potentially to other model users in the United States Air Force and Department of Defense.

Background

The United States Air Force function of sustaining and supporting flight operations in hostile environments, has led to investments in computer simulation models that estimate our capability to operate under combat conditions. These simulations enable decision-makers to evaluate operational concepts and support policies designed to sustain forces employed in combat. They also provide the capability to test concepts and policies and assess the impact of changes without altering the actual system.

In the employment of simulation modeling, the user's confidence in the predictive accuracy of models is, and should be, a key concern. The user accepts the simulation model presumably based on confidence in the model's ability to accurately emulate the real system in question. This confidence in accuracy appears to be directly influenced by verification, validation, credibility and accreditation of the model. Contextual definitions of these terms are examined more thoroughly in Chapter II. In each of the definitions, the simulation model users ultimately accept or reject a model as sufficiently accurate for their purposes, i.e., to support the decision-making process. In an organization as diverse as the United States Air Force it is unlikely that formal development verification and validation activities would produce sufficient confidence in all users that a model meets their needs; this may be the case with AMTAF. Since such efforts have apparently failed to provide widespread user confidence, alternative methods are indicated.

Weapon system simulation models are partitioned into three generally accepted classifications: logistics, airbase, and mission. Logistics simulations provide the means to model support requirements, transportation, and supply processes. The airbase simulations are used to model processes that generate sorties, while mission level simulations provide the tools to estimate sortie effectiveness. To assess the overall capability of a weapon system each of the above environments must be assessed. The Air Force uses several simulation models and decision support tools to evaluate the specific environments, but until recently there was no suite of models that provided an integrated set of simulation models that provided an overall weapon system analysis capability (2:2-1 - 2-3).

AMTAF was developed by Ball Corporation in the mid-1980s, on contract to ASC/XRS. The contracted effort was to develop a simulation model that provided a wider array of simulation capability than the standard models being used (26:1). The AMTAF Sortie Effectiveness Model evolved from the Sortie Air and Ground Engagement Model (SAGE), a weapon system modeling component used widely by the Air Force (2:2-14). As of May 1988, versions of the SAGE model were "resident at AFOTEC (Air Force Operational Test and Evaluation Center), the Brooks Institute, USAF ASC/XRS and XRM (Aeronautical Systems Division Mission Area Planning and Analysis Offices)" (2:2-13). To expand on the acceptance of SAGE AMTAF was designed "to provide the Air Force with a set of tools for evaluating weapon system modifications and designs in support of long-range planning activities" (2:2-14). Ball built upon the SAGE model to produce a four-model suite of capabilities that includes sortie effectiveness, logistics simulation, airbase operability, and threat simulations. These capabilities are partitioned functionally within AMTAF into sortie effectiveness (MASTER), logistics (LOGSIM), sortie generation (SORGEN), and threat (TSARI-

NA) (2:2-12 - 2-14). The final product is an integrated four model simulation package that provides the capability to conduct overall weapon system analysis in each of the environments discussed above: airbase, mission, and logistics (2:2-2).

TSAR was developed by RAND Corporation for the Air Force under the project entitled TSAR/TSARINA. TSAR simulates an environment of theater airbases supported by in-theater transportation, communication, resource management and continental US (CONUS) shipments. Eleven classes of resources are simulated within TSAR, all dealing with airbase operability and sortie generation (18:1). Ball classifies TSAR as a weapon system modeling component used at the airbase level (2:2-12 - 2-14), but as noted above, TSAR's simulation environment extends beyond that of the airbase, e.g., in-theater transportation and CONUS shipments.

TSAR and AMTAF are not fully equivalent, as indicated in the general description of their capabilities. However, of interest in this research are the common airbase simulation or sortie generation features for which the two models are purportedly similar. The TSAR model is accepted and used by the Air Force Center for Studies and Analysis to simulate airbase operability and estimate sortie generation capability. Since AMTAF is said to perform these same basic functions, but is not yet accepted, a direct qualitative and quantitative comparison of the common functional performance of the sortie generation (SORGEN) capability in AMTAF to TSAR provides a basis from which to assess their equivalence.

Justification for the Comparison

Simulation models should be assessed by the using organizations during the development phase of the model's life cycle. Unfortunately it is nearly impossible for all the potential users of a large scale, general purpose model to participate in model development. This fact prompts us to look for alternative methods for model assessment. But what constitutes a sound model assessment methodology? The literature, as portrayed in Chapter II, covers different methods for assessing models using expert opinion, exhaustive analytical means, and real-system data. In some instances one or more of these alternatives is not available. Frequently, the United States Air Force is faced with the inability to collect real-system data because, under some circumstances, its collection would require the destruction of facilities, equipment, and other resources. Under these conditions there is a need for innovative model assessment capabilities. One alternative is the qualitative and quantitative comparison of similar models. Little documented evidence is found that this type of assessment is frequently used. A lack of real system data and the presence of a currently accepted model make this a viable alternative for assessing the level of confidence decision-makers should place in unfamiliar models. This is supported by a 1991 article on simulation assessment procedures by Dr. Saul Gass and several members of the General Accounting Office (GAO). In the article, the authors cite the validation of the Army's 'ADAGE' model against its 'Carmonette' model, and the Air Force validation of 'COMO III' against 'SORTIE' (19:720). The authors go on to say, "The reasonable agreement of results when simulating similar conditions suggests that model-to-model validation can marginally strengthen credibility, especially when comparisons with real-world data are lacking" (19:720).

Problem Statement

The purpose of this study is to present an alternative methodology for comparison of similar models and to demonstrate the methodology by determining the extent to which SORGEN and TSAR are equivalent in terms of sortie generation simulation capability. Equivalence is operationally defined as the similarity of: 1) the level of simulation (mission, airbase, etc), 2) fidelity of simulated functions (comparison of inputs and outputs), 3) ease of use (human interface assessment), and 4) the quantitative statistical similarity of predictions between the models, given equivalent inputs, within a specified confidence level.

Research Objectives

The research is accomplished by developing a methodology to compare the sortie generation capabilities of SORGEN and TSAR simulation models qualitatively and quantitatively, documenting the results, and analyzing the findings.

The five investigative areas are summarized by the following questions. To what extent are the models equivalent with respect to:

1. The general classification and level of performance?
2. The input requirements and characteristics?
3. The output data format and characteristics?
4. The man-machine interface (ease of use)?
5. The output data, given equivalent inputs?

Research Hypotheses

The overall intent of this research is to determine, using the developed methodology, whether SORGEN and TSAR are qualitatively and/or quantitatively equivalent; therefore, the hypotheses posed are:

1. H_O: $SORGEN_{QUALITATIVELY} = TSAR_{QUALITATIVELY}$

   H_A: $SORGEN_{QUALITATIVELY} \neq TSAR_{QUALITATIVELY}$

2. H_O: $SORGEN_{QUANTITATIVELY} = TSAR_{QUANTITATIVELY}$

   H_A: $SORGEN_{QUANTITATIVELY} \neq TSAR_{QUANTITATIVELY}$

Scope and Limitations

The scope of this comparison is to evaluate the simulation models for qualitative and quantitative equivalence in terms of sortie generation at the airbase level. The intent is to compare and exercise, to the greatest extent possible, the models across the full dynamic range of the common functions related to sortie generation. Comparison of the models is limited to a qualitative and quantitative comparison of common features. No attempt is made to evaluate features and capabilities which are not common to both models. This constrains evaluation of the AMTAF model to the SORGEN module. No attempt will be made to evaluate the sortie effectiveness (MASTER), logistics simulation (LOGSIM), or attack (TSARINA) modules of AMTAF. Where the models differ in terms of capability, individual model features are turned off. Where the models differ in terms of input value, but have similar capabilities, every attempt is made to make the databases equivalent. Experimental factors are limited to those applicable to both models except as noted above.

Definition of Terms

Abort - "Failure to accomplish a mission for any reason other than enemy action. It may occur at any point from initiation of operation to destination" (11:1).

Acceptance - The condition that exists when a model user has sufficient confidence in the performance of the model to employ it in the decision-making process.

Accreditation - "An official determination that the model is acceptable for a specific purpose" (32:4).

Aerospace Ground Equipment (AGE) - "All equipment required on the ground to make a weapon system, command and control system, support system, advanced objective, subsystem or end-item of equipment operational in its intended environment" (11:27). Aerospace ground equipment may also be used interchangeably with support equipment.

Air Traffic Control (ATC) - "A service operated by appropriate authority to promote the safe, orderly, and expeditious flow of air traffic" (11:38).

Aircraft Battle Damage Repair (ABDR) - Repair of damage incurred during battle conditions.

Aircraft Ground Damage Repair (AGDR) - Repair of damage incurred during ground handling operations or base attack.

Attrition rate - "A factor, normally expressed as a percentage, reflecting the degree of losses of personnel or nonconsumable supplies due to various causes within a specific period of time" (11:74).

Avionics Intermediate Shops (AIS) - "Special test equipment used for repairing avionic LRUs and SRUs" (15:xi).

Cannibalization - "The authorized removal of specific components from one item of AF property for installation on another item of AF property to meet priority requirements with the obligation of replacing the removed components" (11:107).

Dispersed Operating Base (DOB) - An alternate base of operations where a squadron or wing will deploy, usually with reduced or minimal support capability. Relies on a MOB for extended support requirements.

Line Replaceable Units (LRU) - "An item that is normally removed and replaced as a single unit to correct a deficiency or malfunction on a weapon or support system and item of equipment" (11:393).

LOGSIM - The logistics simulation module of the All Mobile Tactical Air Force Simulation Model.

Main Operating Base (MOB) - A permanent or semi-permanent base of operations normally possessing full maintenance and administrative support capability for one or more operational squadrons. May support dispersed operating bases with aircraft, materiel, and personnel.

Monte-Carlo Simulation - A simulation which samples from a distribution of possible outcomes to obtain a probabilistic approximation for determining the occurrence of events (11:460).

Off-Equipment Task - A maintenance task accomplished on a subassembly or LRU removed from the aircraft.

On-Equipment Task - A maintenance task accomplished on the aircraft.

Scenario - A set of hypothesized conditions derived for the purpose of emulating a specific environment for the purpose of estimating the outcomes of processes.

Shop Replaceable Unit (SRU) - A component part of an aircraft LRU usually replaced at the intermediate shop or depot level of repair. Commonly used to refer to avionics systems subassemblies, such as circuit cards.

Simulation - A descriptive technique that involves developing a model of some real phenomenon and then performing experiments on that model (10:587).

SORGEN - The sortie generation module of the All Mobile Tactical Air Force Simulation Model.

Sortie - "The flight of a single aircraft from takeoff until landing" (11:634).

Tanks, Racks, Adapters, and Pylons (TRAP) - Ancillary aircraft equipment used to configure an aircraft for a specific mission or purpose. Normally managed separately from the weapon system and of key concern to managers due to the importance of the equipment to the employment of the aircraft.

TSARINA - An airbase attack simulation module common to both TSAR and SORGEN.

Validation - "The process of determining the degree to which an model is an accurate representation of the real world from the perspective of the intended uses of the model" (32:4).

Variable - "A characteristic expressed numerically which may differ from one item of observation to another" (11:734).

Verification - "The process of determining that a model implementation accurately represents the developers' conceptual descriptions and specifications" (32:4).

Overview of the Following Chapters

The following chapter reviews pertinent literature relative to simulation models in general, issues of verification, validation, accreditation, and acceptance, and the use of simulation models in logistics decision-making.

The methodology chapter details the development of the techniques used to compare AMTAF and TSAR both in a qualitative and a quantitative sense. Next the measurement techniques, sample size calculations, confidence intervals, decision rationale, and statistical tests are presented, along with the data collection plan detailing measures taken to insure data validity. Necessary assumptions and limitations are included where needed.

The findings and analysis chapter describes the results of both the quantitative and qualitative analytical effort. Comparison tables detailing subjective similarities and differences of the simulation models along with the researchers' impressions of the models' ease of use are also presented here. The statistical comparative analysis of the data resulting from running the models is included with the formal hypothesis statements, confidence level calculations, and power of measurement scores. These data indicate the extent to which the models are equivalent. The findings and analyses are presented and appropriate conclusions are drawn.

The conclusions and recommendations chapter summarizes the research, presents significant findings, and draws conclusions. Finally, the recommended follow-on study areas are presented as an aid to future research.

II. Review of the Literature

Introduction

One of the most prevalent problems with the use of simulation models is establishing their acceptance among the organizations who may benefit from their use. The All Mobile Tactical Air Force (AMTAF) and the Theater Simulation of Airbase Resources (TSAR) models are both airbase operability models. Yet while both models claim to perform the same basic functions, TSAR is used to a much greater extent than AMTAF. The difference in use is indicative of the need for establishing acceptance of simulation models within the organizations intended to use them. This review of literature establishes a knowledge base related to the subject of model verification, validation, and credibility assessment. It provides the ground work necessary for continuing the study of model comparison and acceptance.

Scope of the Research Topic

This review establishes a basic framework from which a detailed study of simulation model acceptance can begin. Definitions are compared of several concepts that are fundamentally important to model acceptance. Some of the existing model verification and validation methods are also explored to examine how they may contribute to acceptance of existing models. Further, a review of examples of model comparisons is made to learn more about the positive and negative aspects of that process.

This literature search begins by defining the subject's most basic elements: verification, validation, credibility, and accreditation. The literature is explored for methods of verification, validation and model comparison. Finally, the literature is summarized and avenues for further research are identified.

Review of Definitions

Four terms are fundamental to the study of simulation model acceptance: verification, validation, credibility, and accreditation. A brief review of the literature reveals that there are no absolute definitions for these terms, but most are basically agreed upon. Each author defines them according to the type of problem being addressed.

Definitions for Verification. In a discussion of verification and validation, Carson defines verification as "the process of comparing the conceptual model with the computer code that implements that conception" (8:552). This definition asks the modeler to determine whether the model's conceptual framework has been successfully captured in the coded instructions. By comparison, Williams and Sikora of the Military Operations Research Society (MORS) have defined verification as "the process of determining that a model implementation accurately represents the developers' conceptual descriptions and specifications" (32:4). In contrast to Carson, Williams and Sikora divide verification into two different subgroups: logical verification, the correctness of equations and algorithms; and code verification, the programming accuracy of the logical elements (32:4). The Law and Kelton definition of verification says that "verification is determining that a simulation computer program performs as intended, i.e., debugging the computer program. Thus, verification checks the translation of the conceptual simulation model (e.g., flowcharts and assumptions) into a correctly working program" (24:299).

Definitions for Validation. Validation is defined by Williams and Sikora as "the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model" (32:4). Validation may be accomplished at one or more levels. The modeler may declare the model's logical description of the real system to be valid, or may declare the model valid for its accurate simulation of real world events. Having achieved validity at either of these levels, the modeler may declare the model valid, but only at that specified level. Williams and Sikora argue that while a model may be validated at many levels, its validity is never absolute (32:4). Law and Kelton provide this definition of validation: "Validation is concerned with determining whether the conceptual simulation model (as opposed to the computer program) is an accurate representation of the system under study" (24:299). They point out that if a model is valid then decisions based on the model will be similar to those made by managers in the real world system (24:299). They noted in earlier writings that accurate representation is determined by comparing the output of the model to the real system (23:334). Shannon's definition of validation stresses the importance of the manager's role. He describes validation as "the process of bringing to an acceptable level the user's confidence that any inference about a system derived from the simulation is correct" (31:29). This definition establishes the need to look more closely at the manager/decision-maker's role in model use and other definitions that describe the manager's confidence in model output.

Definitions of Credibility and Accreditation. While these two terms differ in meaning outside the realm of modeling, they are approximately equal in this context. The General Accounting Office defines credibility as "the level of confidence that a decision-maker should have in their [model's] results" (21:2). Carson states, "a credible model is one that is accepted by the client as being sufficiently accurate to be used as an aid in making decisions" (8:552). Williams and Sikora provide the MORS definition of accreditation to be "an official determination that a model is acceptable for a specific purpose" (32:4). The Law and Kelton definition of credibility is "when a simulation model and its results are accepted by the manager/client as being valid, and are used as an aid in making decisions" (24:299). In each of these definitions, a firm's chosen representative accepts a model as sufficiently accurate to be used for decision-making. The manager's acceptance that a model's output is valid may be as important as the validation process itself, since without the decision-maker's trust that the model will return useable information, there is little chance the model's output will be used and thus little reason for the existence of the model. Given an established accuracy, what methods should be used to establish acceptance in the eyes of the decision-maker?

Review of Methods for Model Acceptance

As expected, the literature revealed several different concepts for assuring the acceptance of models. The most common theme among these concepts is that the most appropriate way to assure a model's acceptance is to conduct verification and validation studies during the development of the model (1:2, 8:552, 13, 23:44).

Furthermore, to assure acceptance of a model, the developer should include the intended users in every step of the development process. Participation during the development of the model allows the user to determine personally that the model does what it is intended to do (8:552). Failure to follow this practice causes a failure of confidence, as evidenced by the case of AMTAF, a fully developed Air Force weapon system modeling simulation. The organization that procured AMTAF was involved in the development process but the organization responsible for performing analysis was not and consequently lacks the confidence necessary to use its output. In some cases, as with AMTAF, the end user is not able to establish model confidence as a result of direct participation in model development. This being the case, alternative methods should be sought for helping the user achieve a sufficient level of confidence (acceptance) in their models. The remainder of this literature review examines the existing concepts of verification, validation, and credibility assessment and model comparisons to determine a baseline from which post-development acceptance may be achieved.

The Balci Method. Osman Balci developed a concept of the simulation life cycle, presented in Figure 1. His method includes 10 data input/output phases (rectangles), 10 data transformation processes (hollow arrowheads), and 13 credibility assessment stages (CAS) (solid arrowheads) that come together to form his perception of the simulation development life cycle (1:62). Of primary interest in this research are the 13 credibility assessment stages (Figure 2) that Balci developed (1:66), which illustrate a wide array of validation and verification requirements. Balci's credibility assessment stages provide a framework of the elements involved in validation and verification, and the relationships that may exists among them.

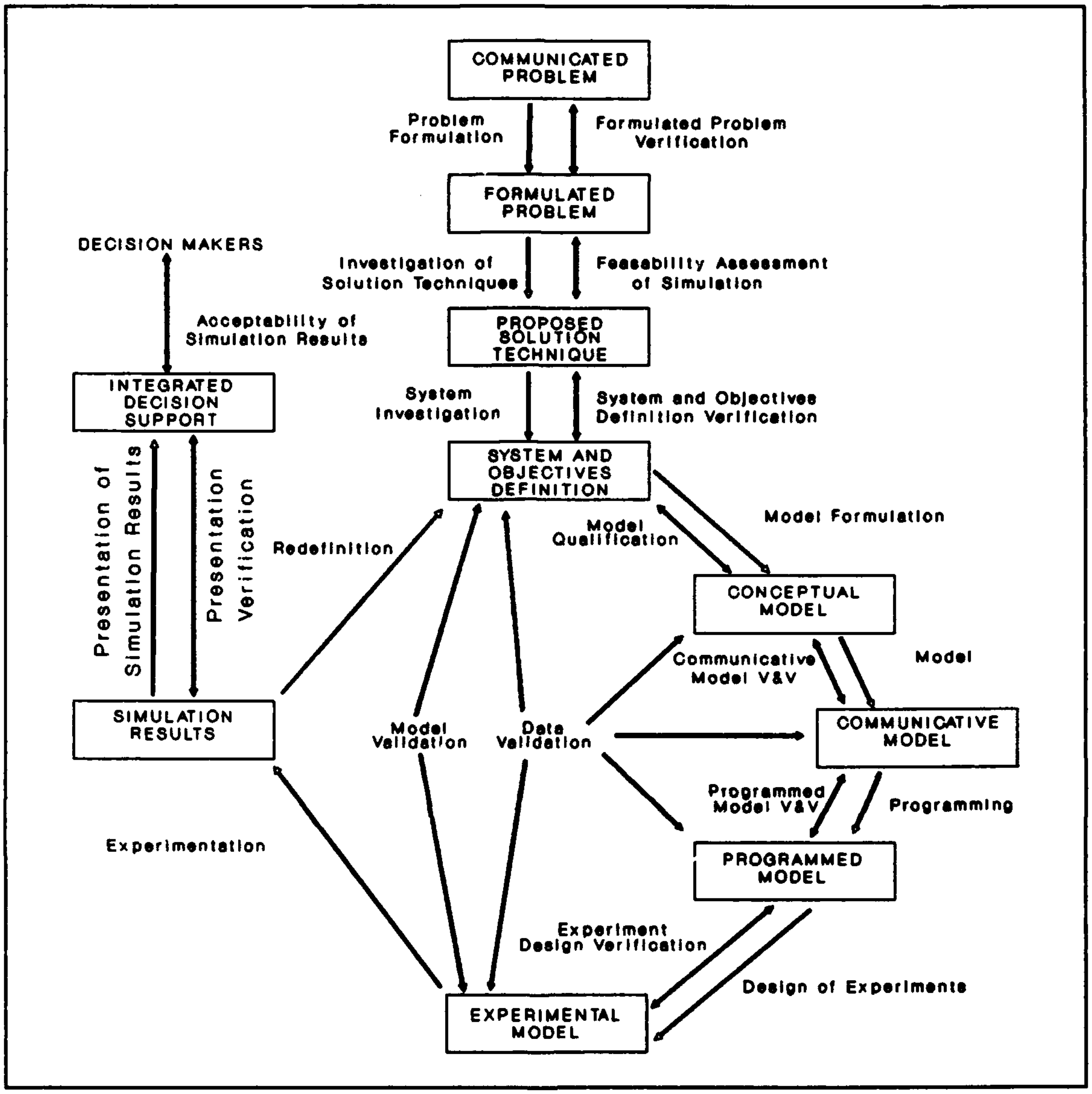

Figure 1. The Balci Life Cycle (1:63)

Balci intended this assessment framework to be used during the development stages of modeling. The problem being experienced by the Air Force is concerned with fully developed models, requiring modification of Balci's work before its use. However, Balci provides both a logical and quantitative framework for determining which elements are most important in assessing the credibility of an existing model.

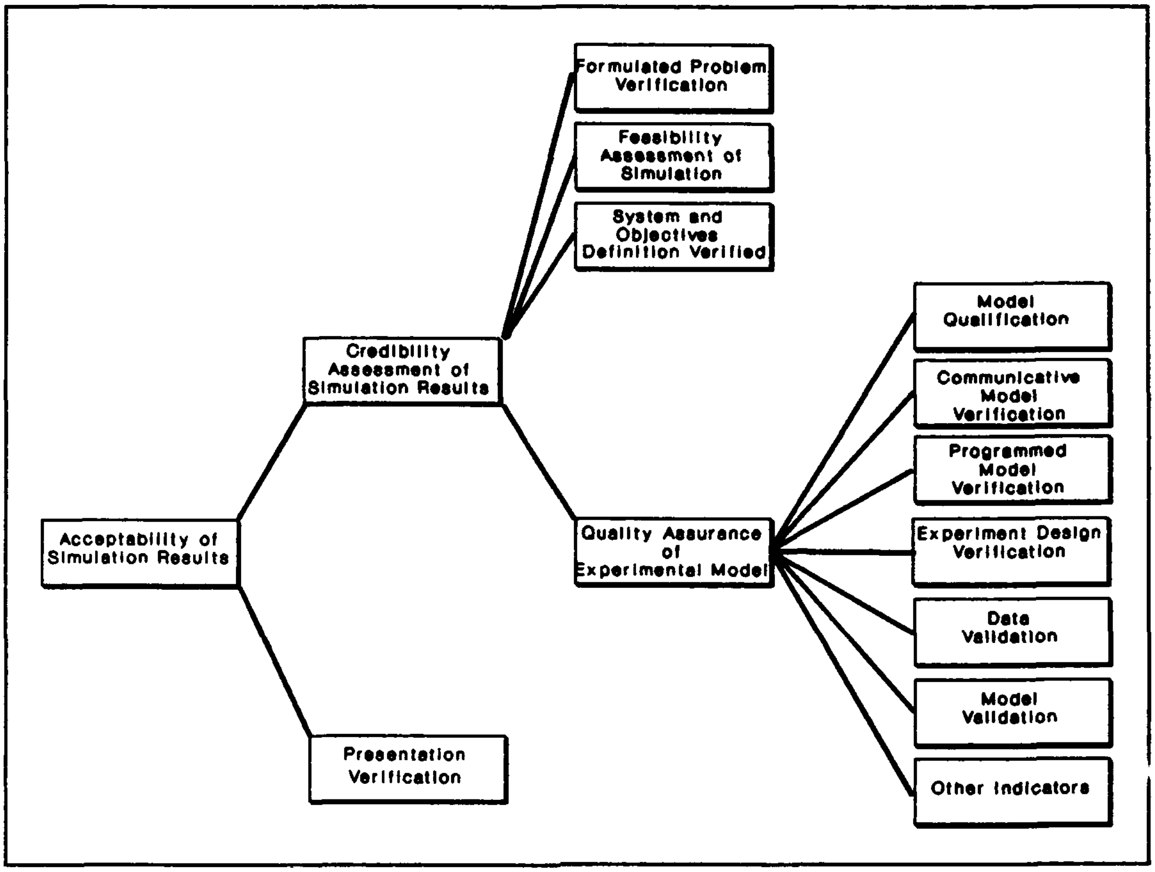

Figure 2. The Balci Credibility Assessment Stages (1:66)

The Williams and Sikora Method. Williams and Sikora provide a more straight forward approach to model accreditation. They divide the tasks associated with model accreditation into two major segments: face validity and documentation (32:5).

Face Validity. Face validity is accomplished through a comprehensive review of the model output data by comparing them to data collected from the real system. This type of validity is frequently performed using experts to evaluate the available data. Expert evaluation of models is very useful especially where real-world data are not available (32:5). Law and Kelton refer to a model that has high face validity as one that "on the surface, seems reasonable to people who are knowledgeable about the system under study" (24:308). Law and Kelton's discussion of establishing face validity recommends the modeler use at least seven techniques. First, obtain information from system experts by working closely with individuals who are knowledgeable of the system under study. Second, collect information about the system by collecting information from an identical or similar system as the one of interest. Third, make use of existing theory when at all possible. Fourth, make use of the results from similar modeling efforts which provide a good source for lessons learned. Fifth, the modeler should make use of experience and intuition, being careful to substantiate these hypotheses during the modeling effort. Sixth, bring the decision-makers and managers into the modeling effort on a regular basis, allowing them to develop an understanding of the model. This activity will help assure the decision-makers develop a trust in the model, its capability, and its output, making the use of the model for decision-making more likely. Finally, the modeler should conduct a formal walkthrough of the conceptual model with managers, decision-makers, and other key personnel. The walkthrough assures key personnel that the model's concept is sound and that assumptions are correct (24:308-310). These seven steps provide a sound approach for assuring the model's logical design is adequate for its intended purposes.

Documentation. The second set of tasks suggested by Williams and Sikora, documentation, is intended to cover both logical verification and code verification. Logical verification covers assumptions and the review of pathways through the model. Code verification takes a look at the actual programmed code to ensure it follows the structure of the conceptual model and correctly addresses the assumptions of the model. Documentation is used heavily by the analyst to perform the logical verification of the model and is also used in conjunction with the programmed code to perform the code verification (32:5-6).

The General Accounting Office Methods. The General Accounting Office (GAO) is credited here with the development of two methods for assessing the credibility of simulation models. The first was published in 1979 and considered five criteria, while the second was published in 1987 and covered three broad categories.

General Accounting Office, 1979. The 1979 GAO publication, "Guidelines for Model Evaluation," provides an alternative for model evaluation. This method of evaluation uses five major criteria: 1) documentation, 2) validity (theoretical, data, and operational), 3) computer model verification, 4) maintainability (updating and review), and 5) useability (Figure 3). This method's foundation is based on the model's documentation and the evaluation of the model is based on validity, verification, useability, and maintainability (20:3, 22:9).

General Accounting Office, 1987. The General Accounting Office (GAO), in a 1987 report on simulation assessment procedures, proposed a framework for assessing the credibility of Department of Defense (DOD) simulation models. The framework consists of three major areas of concern: 1) Theory, model design, and input data; 2) Correspondence between the model and the real world; and 3) Support structures, documentation, and reporting (19:713-714, 21:19). These elements are considered by the GAO to constitute important elements in assessing the credibility of models.

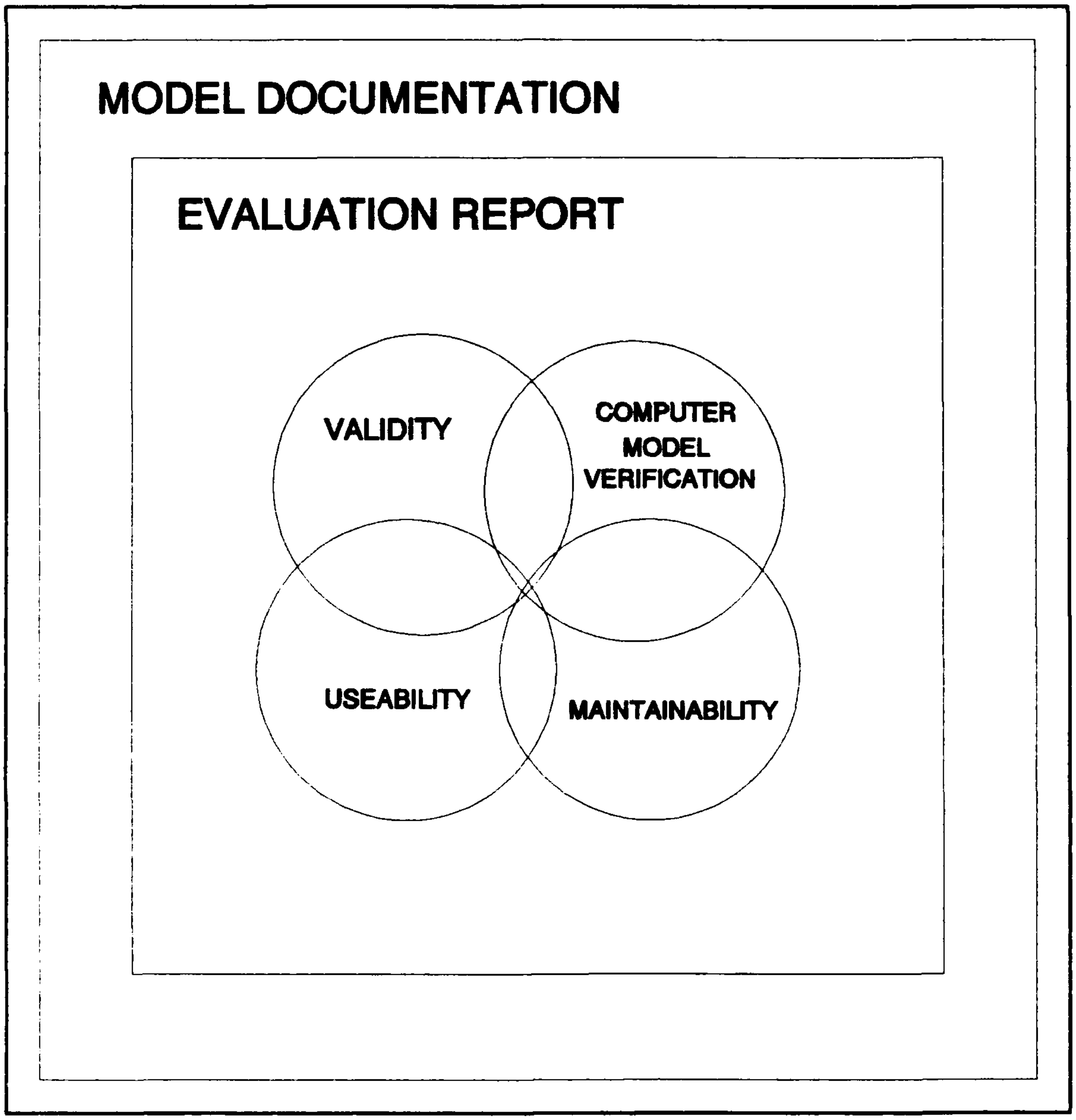

Figure 3. Interrelationships Among Evaluation Criteria (22:25)

Theory, Model Design, and Input Data. This division of the GAO framework is concerned with how well the model simulates the real system. It considers such things as the characterization of the real world used to develop the model, whether the model's conceptual framework matches the real world characterization, and how accurate and realistic input data are compared with those found in the real system (19:716-718, 21:18-20).

Model Versus Real World Correspondence. This area is quantitative. It addresses accuracy of verification efforts, statistical accuracy, validation accuracy, and sensitivity testing. The objective of this area is to demonstrate that the model provides an adequate representation of the real system and that the data produced by the model are consistent with data produced by the real system (19:718-720, 21:20-21).

Support, Documentation, and Reporting. Their last area addresses issues of infrastructure dealing with the management of model support requirements. Some of the issues covered in this portion of the credibility assessment are: 1) ensuring accurate documentation is available for personnel who use the model, 2) ensuring requirements for design, operations and data management are in place and functioning, and 3) ensuring accurate reports are provided concerning the operation of the model (19:720-721, 21:21).

The Three Step Method. Carson, Law and Kelton, and Van Horn (8:552-558, 24:307-314, and 31:247-258) all use the same three steps as the framework for their methods of assessment: 1) face validity, 2) assumption testing, and 3) output testing. Each of these three steps is discussed individually in the following sections.

Testing Face Validity. Face validity is a test of reasonableness from the perspective of the expert user. The model should be reviewed to ensure that the conceptual model matches the real system to an acceptable degree. Techniques that might be employed here are: tracing logical paths through the model to ensure conflicts are removed, examination of the methods used for structured programming, and conducting sensitivity analysis to determine which variables are most sensitive to input changes (8:554-555). Law and Kelton state that evidence of expert involvement in the development of models improves the face validity of the model (24:308-310).

Testing the Assumptions. In their discussions of testing assumptions, both Van Horn, and Law and Kelton support the use of empirical analysis. Their primary tool for accomplishing these tests is sensitivity analysis, the analysis of models to determine what changes in output data occur when model variables are experimentally tested (24:310-311, 31:251-252). Carson argues that the assumptions can be tested through animation. The use of graphical presentations enhances the expert's, as well as the user's, confidence that the model assumptions are correct. Carson warns, however, against making judgements concerning the accuracy of assumptions based on animation alone and adds that statistical analysis should complement this activity (8:555).

Testing Output Data. Law and Kelton provide a basis for beginning the test of output data. They discuss the Turing test which consists of providing experts with output data from the real system as well as the model. They are asked to identify which set of data came from the real system. If the experts are not able to distinguish between the two sets of data, then the model probably approximates the real system with some degree of accuracy. They also discuss the use of statistical analysis, primarily whether one should use hypothesis testing or statistical differences. Since the model is only an approximation of the real system, the use of a null hypothesis that says the real system and the model are equal is logically wrong. Law and Kelton support the use of difference testing to determine whether there is a statistical difference between the output of the model and the real system (23:340-342).

Review of Model Comparisons

Through review of the literature it is evident that the desired method of evaluation is to involve prospective users in the development of simulation models, thereby enabling the user to simultaneously develop a sense of confidence in the structure and capabilities of the model. It is also desirable to evaluate models using the best available real data, preferably complete historical data, that would allow the evaluation team to observe whether the model accurately replicates historical results. Unfortunately, some instances render these conditions unobtainable.

In large organizations such as the United States Air Force, it would be impossible to include every potential user of a model in its development. Furthermore, confidence in a model must be established by the user for a particular purpose (20:7). There can be no blanket evaluation of a model that applies to every situation. Many of the airbase operability (ABO) models used by the Department of Defense simulate the effect of air and ground assaults on facilities, equipment, personnel, and other resources of airbases. To collect and use real data would require the destruction of these resources; this is obviously unacceptable. These situations demand the development and use of alternative methods for evaluating simulation models.

The literature reveals other alternatives for simulation evaluation such as the use of expert opinion and model-to-model comparisons. While these and other evaluation methodologies have been studied at length, it is model-to-model comparison that may hold considerable value and deserves further research. Within this realm of model evaluation, there appears to be little evidence in the literature that model-to-model comparisons have received in-depth attention.

There are examples of model comparisons in the literature. In September 1986, an AFIT thesis authored by David Noble, titled "Comparison of the TSAR Model to the LCOM Model," attempted to compare two models. Noble's effort consisted of a statistical analysis of the two models using a randomized block design and manipulating one independent variable, target daily sortie rate, while holding all other independent variables constant. He then observed the two dependent variables, man-hour usage and sortie production, and analyzed the output of the two models to determine if significant statistical differences existed between the variable means (29:12-20). Noble concluded his experimental design was flawed due to differences between the databases used for each model (29:12-20).

A similar effort in September 1987, by Gregg Clark, titled "The Theater Simulation of Airbase Resources and Logistics Composite Models: A Comparison," also attempted to compare one model to another. Clark mirrored the work of Noble but attempted to ensure the databases were as identical as possible (9:24-33). Clark concluded there were no significant differences in the results of the two models, this seemingly due to nearly identical databases (9:58-59).

An additional research effort was conducted in 1991, by David Leonhardt, titled "A Comparison of the All Mobile Tactical Air Force and Logistics Composite Simulation Models." In his comparison Leonhardt concluded there was no statistical difference between the two models under study (26). However, Leonhardt like his two predecessors only varied two factors to achieve his results, and the databases used were very small and general.

It should be pointed out that although Clark's and Leonhardt's study showed no significant difference in the output of the two models, the experimental designs used by Noble, Clark, and Leonhardt were very tightly controlled and limited in scope. In all three studies the researchers only marginally exercised the quantitative capability of the models and made no documented comparison of the models' features. A qualitative and quantitative comparison is required for a sound model assessment.

In a 1979 report by the GAO on DOD simulations, the authors discuss both the Army's and the Air Forces's use of model-to-model comparison for the purposes of validation. They note, "The reasonable agreement of results when simulating similar conditions suggests that model-to-model validation can marginally strengthen credibility, especially when comparisons with real-world data are lacking" (19:720, 21:45). This approach to model evaluation seems reasonable, but requires a well- documented, substantive methodology. This research addresses the development and documentation of that methodology.

Conclusion

This literature review provides the basis for beginning an in-depth study of acceptance as it pertains to fully developed models. The definitions discussed provide a foundation and common point of reference for the research. The methods discussed offer a list of steps that can be taken to achieve acceptance of simulation models. The execution of these methods, however, is intended to occur during the development stages of modeling. They differ from the model assessment method developed in this research because the method developed in this research takes place in the post-development phase of the model's life-cycle.

III. Methodology

Introduction

The literature review outlined methods of verification and validation used during model development to ensure the accuracy of simulation models. The cited methods are formal processes that provide the basis for initial acceptance. Frequently model users are not exposed to the formal verification and validation process, as are modelers, leaving them to develop their own methods for establishing a basis for model acceptance after the development phase. In the absence of a formalized methodology model users do one of three things: research and establish a basis for acceptance or rejection, fail to use the model from lack of confidence in its ability to support management decisions, or use the model without properly questioning its validity, the most dangerous of the three options. There is little written regarding a means of insuring post-development accuracy of models short of comparing model data to some accepted standard of measurement, usually real system data as in the Turing test. In some instances, however, collection of real world data is impractical or impossible; this is especially true for airbase operability and attack data. Model-to-model comparison offers a potential solution for documenting a model's capability, qualitatively comparing its purpose and features, and quantitatively testing its performance, enabling the user to select the model that best meets his or her needs. Further, the understanding and insights resulting from such a comparison establish a foundation on which to base a level of confidence that the model effectively supports the decision-making process. Earlier research in model comparison has focused solely on the quantitative statistical equivalence of models. The Noble (29) and Clark (9) theses established some comparative equivalence of portions of the Logistics Composite Model (LCOM) and TSAR, followed by the work of Leonhardt (26) who conducted similar comparisons with LCOM and AMTAF. A more robust quantitative comparison and the addition of a qualitative comparison is necessary to fully assess overall model equivalence.

The foundation of this research is adapted from a construct used by the GAO, discussed and presented as Figure 3 in the literature review. Figure 3 is repeated below as Figure 4 to ease its comparison to the framework for this research which is presented in Figure 5. Note the common basis of the two diagrams: model documentation. The design, programmer, analyst, and user manuals are key to the evaluation of any model. The proposed evaluation and research efforts rely on the depth and accuracy of the model documentation to support the comparison of AMTAF and TSAR. Since comparison of existing models is necessarily done after the model development phase has ended, it is extremely important that all available documentation be acquired and studied in depth to fully comprehend the intended use, features, and operation of the models under study. The research begins by examining the models from a qualitative perspective.

Qualitative Comparison

The qualitative evaluation of a simulation model is necessary to provide the potential user the requisite level of knowledge and understanding to competently employ the model and its capability, and answers the first four investigative questions.

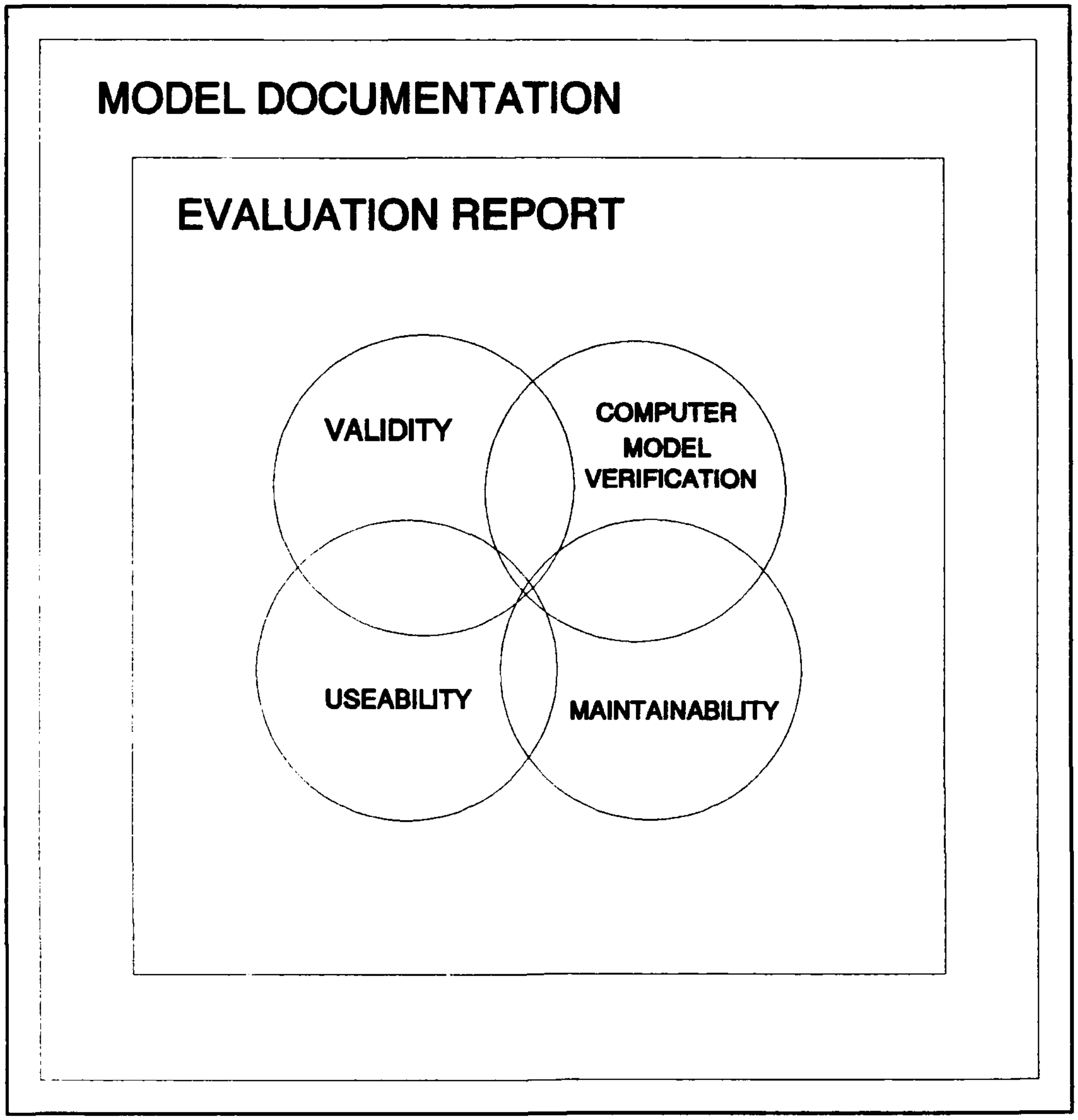

Figure 4. Interrelationships Among Evaluation Criteria (22:25)

To what extent are the models equivalent with respect to:

1. The general classification and level of performance?
2. The input requirements and characteristics?
3. The output data format and characteristics?
4. The man-machine interface (ease of use)?

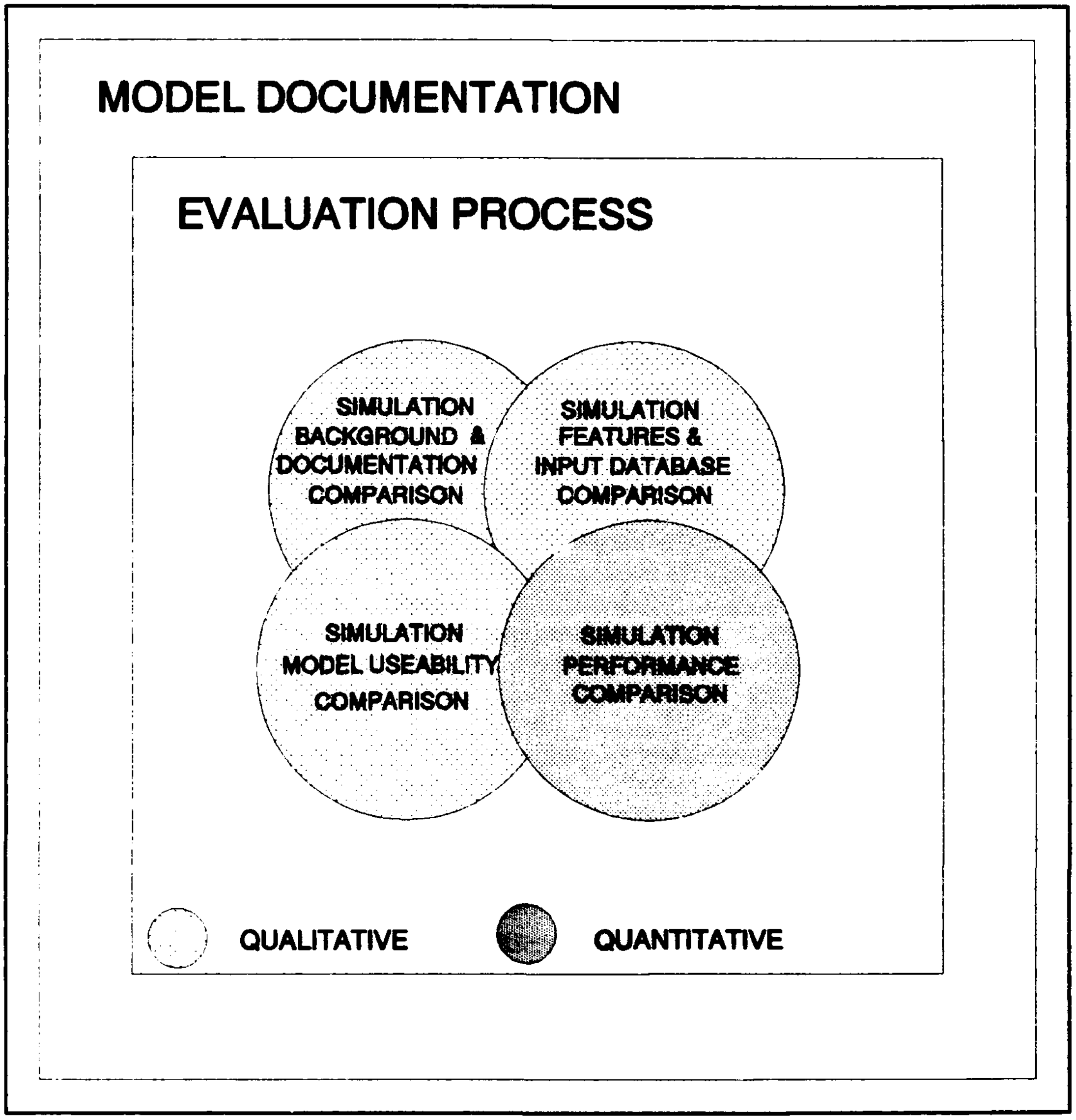

Figure 5. Construct For Model Comparison

Each of these questions addresses an area within the comparison construct presented in Figure 5. The first question is addressed by the simulation background and documentation comparison. The second question is addressed by the simulation features and input database comparison. Finally, the third and fourth questions are addressed within the simulation model useability comparison, since output data and the man-machine interface are factors of useability.

The potential user can establish a basic knowledge of the models under study by conducting an in-depth review of the model's documentation. This section outlines the qualitative comparison methodology used to achieve a detailed understanding of TSAR and SORGEN. An explanation is given why a thorough review of the documentation is important, including the models' general classification, level of performance, modeling environment, evolution, and its original reason for development. Next, the documentation of each model is studied to determine whether both models' documentation is sufficient for the purposes of future modeling efforts. During the documentation comparison the features of both models are tabulated in a form that permits easy comparison of capabilities. The models' data input requirements are also compared through the construction of a logical map of the models' databases, and an estimate of the models' useability is documented based on our own subjective criteria. The qualitative evaluation begins with the background and documentation of the models which is shown in one of the circles of Figure 5. Figure 6 provides an expansion of the model background and documentation comparison.

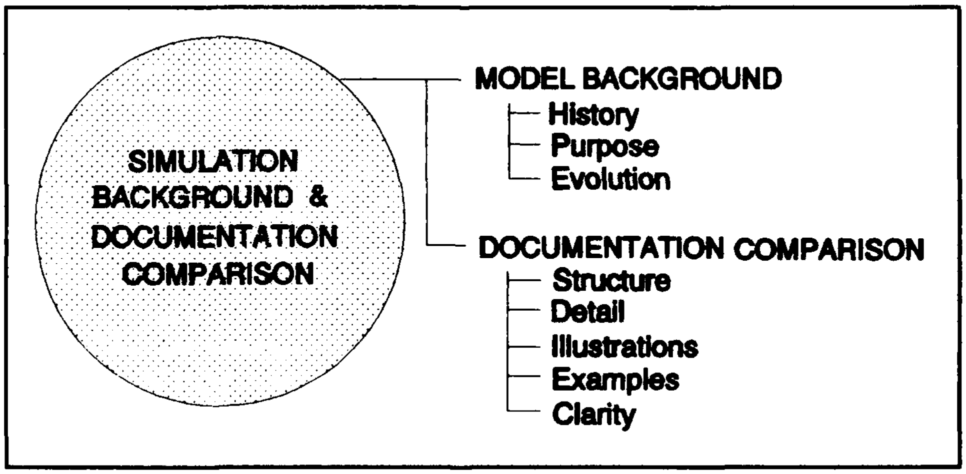

Figure 6. Background and Documentation Comparison

Model Background. A basic knowledge of each models' evolution is necessary to understanding what the models were originally intended to simulate. The researchers review the history of the models to determine when the models were originally developed, their original purposes for development and their evolution since development. The documentation for both SORGEN and TSAR is researched to learn the general classification of the models (e.g. sortie generation, attack, logistics, etc), and the level of performance (e.g. high, intermediate, or low fidelity). These findings are then documented and provide a general description of each model under study that may be referred to in future study or by other researchers or model users. The next phase is comprised of comparing the documentation for TSAR and SORGEN.

Documentation Comparison. Once the basic comparison of the models' classification has been accomplished, the specific features, capabilities, and characteristics are investigated using the documentation of each model as the basis of this analysis (20:3, 22:9). This investigation also compares the models' documentation and assesses the relative strengths and weaknesses along with similarities and differences.

The documentation is evaluated holistically, considering structure, level of detail, illustrations (diagrams & figures), examples, and ease of comprehension (clarity). This subjective analysis is drawn from the combined perceptions of the researchers. Since neither investigator has prior experience with either simulation model this analysis should provide a reasonably unbiased measure of the adequacy and useability of the documentation. The strengths and weaknesses of each set of documentation are recorded and then compared and contrasted to establish the degree of similarity. The overall intent of this analysis is to provide a measure of the suitability of the documentation to support employment of the models. Continuing on to the next qualitative area shown in Figure 5, an assessment is made of the models' features and databases. This portion of the qualitative comparison is expanded in Figure 7.

Features Catalog and Database Comparison. To compare the simulation models' specific features, capabilities, and characteristics, a catalog or listing of the features is assembled relying on the simulation models' written documentation to gain knowledge and insight into the models' individual capabilities. Comparison of the models' databases is accomplished by identifying common data requirements and mapping the data locations within each of the respective databases. Identifying the purpose and location of each data element provides the basis for the translation of a database for use in quantitatively comparing the two simulation models. The construction of a table of features is accomplished first.

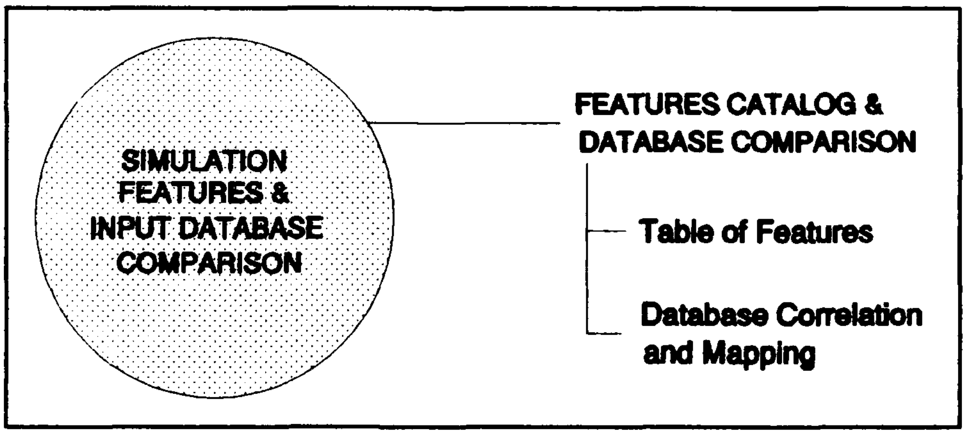

Figure 7. Features and Database Comparison

Table of Features. A list of the models' capabilities is constructed and sorted alphabetically to produce a comprehensive listing of simulation features. These features are placed into a table the format of which is presented in Table 1. Assigned to each feature is a yes/no indication of its presence or absence in each model. The table provides a tool from which to compare the capabilities of the two simulation models. In those areas where a simple yes or no is less than adequate in describing the presence or absence of a feature, exceptions are indicated and footnoted specifically within the context of the table. Once completed the attention is turned to the database correlation and mapping effort.

Table 1

Simulation Feature Table Format

Simulation Feature	TSAR	SORGEN

Database Correlation and Mapping. The TSAR and AMTAF simulation models are driven by databases that provide information describing base facilities, personnel, aircraft, maintenance, supplies, policy and so on. Any similarity in database structure and content should indicate some degree of commonality between the simulation models. Since the goal of this research is to determine the extent to which the SORGEN module of AMTAF is equivalent to TSAR, it is necessary to construct databases that are equivalent to facilitate the quantitative comparison of the models. Since TSAR is the more widely accepted of the two simulation models, a representative TSAR database is translated to one useable by SORGEN. Identification of the required data elements and the determination of where they reside in the respective databases is critical to the translation process. To accommodate the database translation process and to expand the qualitative comparison of the models the two databases are mapped or cross-indexed.

The TSAR database is organized in fixed, 80-column IBM card format. It is loosely organized by card type, and not readable by most users without indexing information. The AMTAF data are organized into functional databases containing single or multiple relations. These relations consist of functionally related data elements. The AMTAF database printouts label the data elements for readability.

The AMTAF database structure is studied first to determine what data are required by the SORGEN module. When the data are fully identified, a search for correlative data within the TSAR database is accomplished. A listing of the required AMTAF databases is used as a basis of this effort to take advantage of the data labeling feature. The location of identical or similar data in the TSAR database is cross-indexed using the card type (CT) number and card column indicators. Database mapping tools are formulated that permit the cross-indexing of the AMTAF database structure to TSAR card type and alternatively by TSAR card type to AMTAF database structure. These mapping tools provide the key for use in the database translation process. The final category of the qualitative comparison from Figure 5 is the assessment of the models' useability. Figure 8 provides an expansion of model useability as it relates to the model comparison.

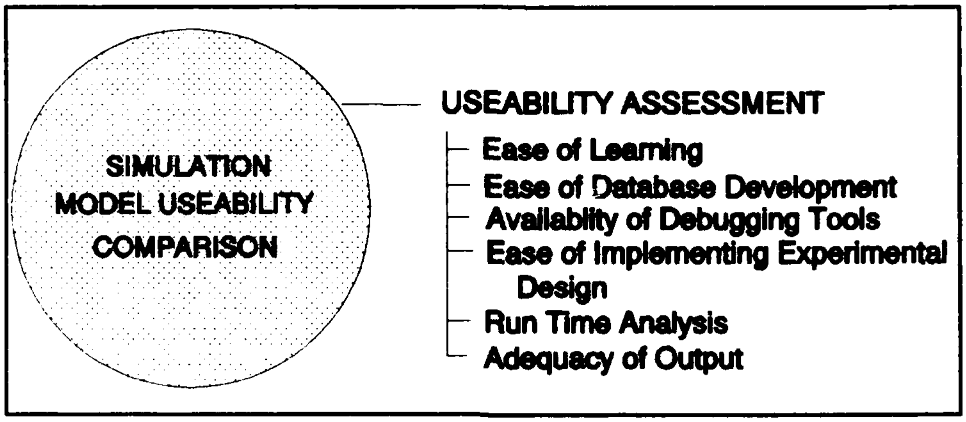

Figure 8. Useability Comparison

Useability Assessment. An assessment of the useability, or ease with which a model is employed, may depend on several factors, among them the users' past experience in modeling, preferences for menu driven versus non-menu driven software, and the depth of study prior to the first modeling attempt. Since useability is different for individual users, it is appropriate to conduct a subjective evaluation of model useability over the course of the entire research project. The useability assessment is necessarily longitudinal in nature and the model user should be conscious of its importance from the beginning of the model comparison. Six areas of useability believed important to the success of the simulation effort are subjectively assessed. First, an assessment is made of the differences that exist for each model in terms of the difficulty in learning to run the models. Second, an assessment is made of the problems encountered during the production of databases for each model and also the difficulty experienced with manipulating the needed simulation data reports. Next, an assessment is conducted for the availability of debugging tools for each model and the ease with which these can be used to correct problems. The fourth assessment involves identifying the problems encountered in implementing the experimental design, which includes the ease with which the model accommodates the manipulation of input data to achieve a successful experimental design, and whether multiple simulation runs may be submitted concurrently. Fifth, an assessment is made of the time required to achieve each run, if batch submissions may be made, and whether the user's presence is necessary throughout a run. The final assessment is for the adequacy of the simulations output, the content and format of system generated output, whether the user is provided with an ability to define output reports to suit specific needs, and whether the simulations provide the type of data necessary for the user to make informed decisions based on the results.

Useability is an important factor in the assessment of a model, especially from the user's point of view. It is useability that may have the most significant effect on the user's attitude toward a model and thus determine whether a model is used for decision-making purposes. The qualitative comparison of models is necessary for determining whether there is a logical basis on which to make an overall comparison of models. Completion of the qualitative comparison prepares the way for building common databases and conducting a quantitative comparison of the models. The quantitative comparison is expanded and presented graphically in Figure 9.

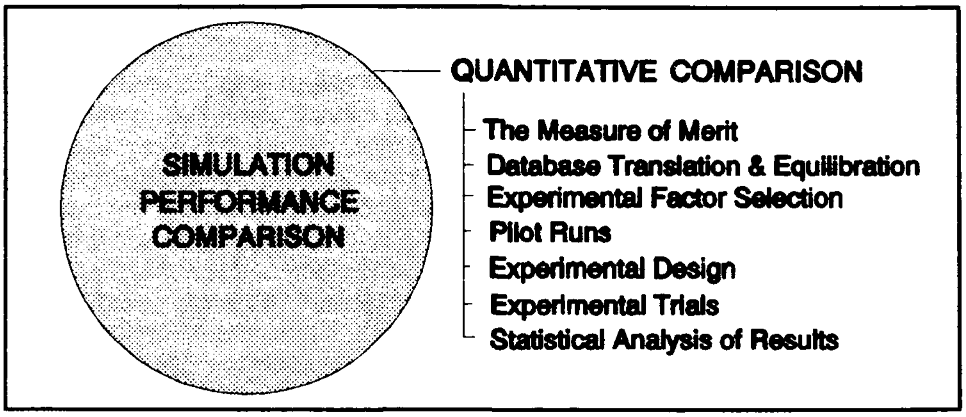

Figure 9. Quantitative Comparison

Quantitative Comparison

The quantitative comparison of simulation output provides the model user with another class of information on which to evaluate the equivalency of the models under study and answers the investigative question: Are the model outputs equivalent given equivalent inputs? In the literature review an investigation is made into several methods by which researchers can assess the accuracy of model output; these are adapted and synthesized for the comparison of TSAR and SORGEN. The quantitative comparison of model results begins with a determination of the important measures to be used by the researchers. The translation of the databases is discussed and how they are equilibrated to the greatest extent possible. A discussion is also provided of how the factors are chosen for the experimental design. A methodology is provided for conducting pilot runs, determining the actual experimental design, running the experimental trials, and conducting the statistical analysis. These methodologies constitute the quantitative portion of the research and are discussed in detail in the forthcoming sections of this chapter. The discussion begins with the measure of merit.

The Measure of Merit. Conducting a quantitative comparison requires common numerical measures of overall performance. If possible the effect of all aspects of the simulation environment should be observed in a single measure. Since the TSAR and SORGEN models are designed to simulate the operation of a military airbase, sorties generated, i.e. the number of combat missions flown during the simulation, provides an overall measure of airbase operations which is meaningful to the purpose of this research. Assuming the simulation models capture the interactive airbase functions which influence overall sortie production, sorties generated will be the measure of merit for the purpose of quantitatively comparing TSAR and SORGEN. To conduct this comparison a TSAR database is translated for use in SORGEN and the two databases are equilibrated to employ common model features.

Database Translation and Equilibration. Quantitatively comparing TSAR and SORGEN requires a common input useable by both models. Much of the research conducted in the qualitative analysis contributes directly to this activity. The development of the table of features to permit comparing the models from a functional perspective and the preparation of the database mapping tools provide the mechanisms to support the development of equivalent input databases.

There are several databases available for TSAR that have been developed for the purpose of establishing policy, determining manning and resource levels for theater operations, and for studying the effect of enemy attacks on airbases. The F-15 TSAR database has been thoroughly tested and used previously in extensive research. It was originally developed by Milt Kamins for RAND in 1987, based on a data collection effort by RAND, AFCSA/SAGP, and SYNERGY. Therefore, the F-15 database was chosen as a proven and acceptable basis from which to develop equivalent databases.

Dissimilarities in the simulation models and their databases should be no surprise since different modelers developed them. When a feature is encountered that is not in both models or is not obviously replicated equally, the lesser capability becomes the standard and the model possessing extended capability in that particular function is constrained to a level that, as nearly as possible, equates it with the lesser capability. Where differences in fidelity (level of detail) are encountered, the data inputs are aggregated or disaggregated, where possible, to exercise the maximum number of common functions. These steps are designed to exercise TSAR and SORGEN as thoroughly as this research will permit.

After completing the database development process, each database is thoroughly tested in its companion simulation model using the debugging tools and error detection features available. Once the databases are proven operable, they are used to create the various versions needed to experimentally exercise the simulation models. First, the factors are chosen to be used in the comparison experiment.

Experimental Factor Selection. Several factors important to the sortie generation process must be chosen to build the experimental design for this research. These factors must be clearly defined in the input databases and capable of being varied to facilitate experimentation. In addition, the following criteria are used in the selection of factors:

1) Each factor must be present in each of the simulation models or capable of being implemented with a high degree of equality.

2) Each factor must have a direct impact on sortie generation and be related to the logistics infrastructure.

3) Each factor must be directly influenced by operational demands, i.e., increased sortie rates, or use of more logistics resources.

4) The set of factors establishes a broad inference space, i.e., a significant portion of the logistics infrastructure is encompassed by the factors chosen.

Each of the models under study simulates the logistics operations of combat-oriented airbases, with the ultimate goal of flying combat aircraft missions. The factors chosen should represent the major capabilities of the logistics operation, i.e. fuel, munitions, support equipment, etc. A general representation of the combat-oriented logistics system and its components is shown in Figure 10. In the Noble (29), Clark (9), and Leonhardt (26) research, only one or two factors are varied in the experimental design. To achieve a more dynamic and realistic simulation of an airbase's capability to generate sorties, multiple factors must be exercised. This also allows a broader inference space on which to evaluate similarity of model performance.

Preliminary study by the researchers show that the two simulation models have a high degree of commonality in input data. Review of earlier research done on the TSAR model by Diener (12) identified factors which could be used in this research. These efforts narrowed the focus to eight factors which meet the above criteria; however, database development may require altering the choice of factors. The tentative factors and their lowercase letter assignments are:

1. Aircraft - factor a
2. POL (fuel) - factor b
3. Munitions - factor c
4. Missions - factor d
5. Personnel - factor e
6. Spares (parts) - factor f
7. AIS (avionics intermediate shops) - factor g
8. Support equipment - factor h

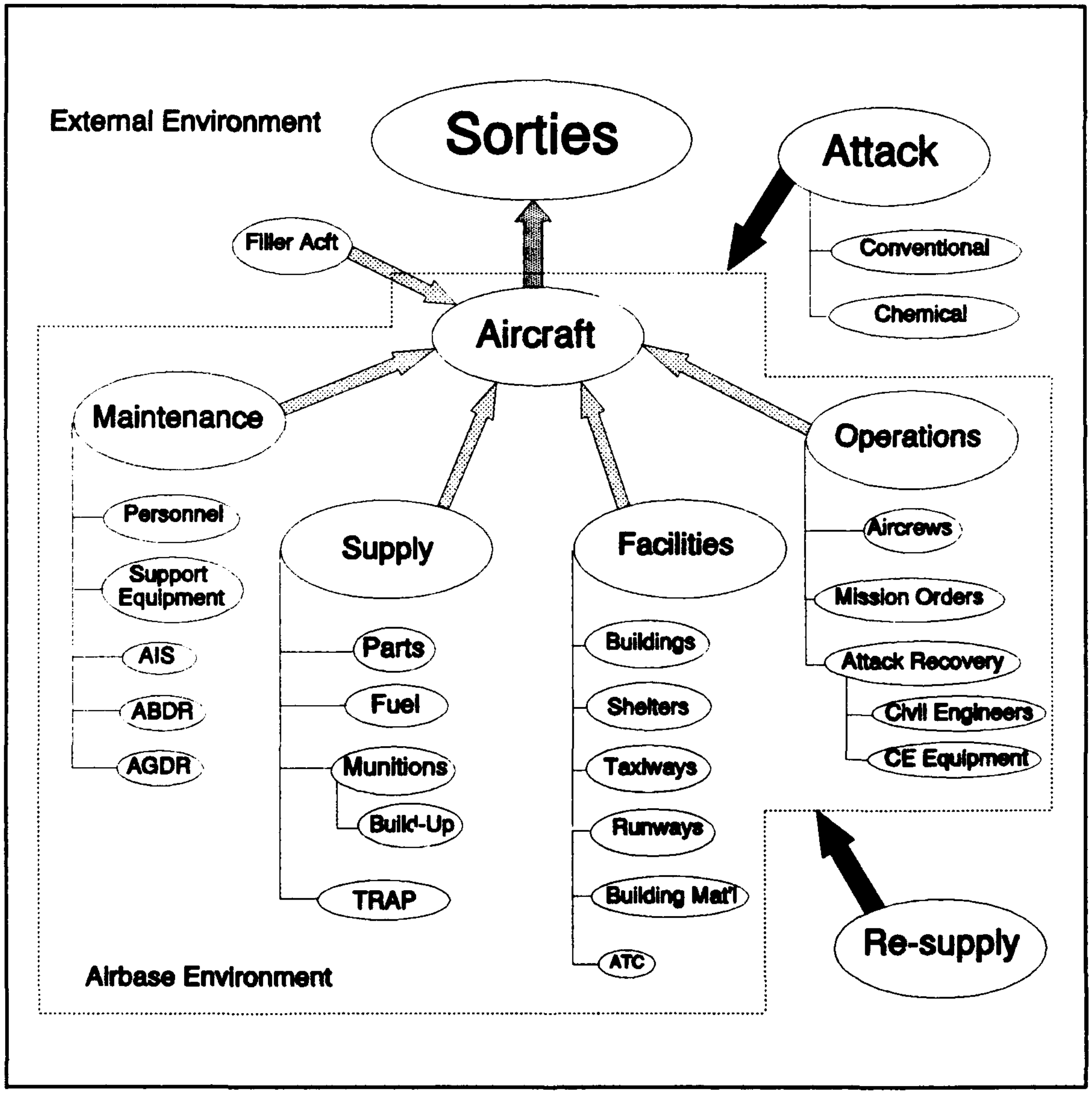

Figure 10. Sortie Generation Logistics Environment

Selection of factor levels must be accomplished next. Drawing again on previous research, factors a,e,f,g, and h have established high and low treatment levels which are drawn directly into this research (12). The levels used in earlier research are, for the high levels, indicative of real world values based on the experience of the researcher. The low-level values are set at a 25% reduction of the high-level values. Factors b (POL) and c (munitions), also used in Diener's research (12), are changed to represent different quantities of resupply. The high-level resupply provides 90% of the POL and munitions required if all of the high-level mission demands are achieved. The low-level resupply provides 75% of the POL and munitions required if all of the high-level mission demands are achieved. Factor d (missions) is unique to this experimental design, and is set via a sortie demand schedule. The value initially selected for the high level establishes a maximum sortie demand of 4.7 sorties per aircraft per day. The low level establishes a maximum sortie demand of 2.0 sorties per aircraft per day. The factor levels are then evaluated using a series of pilot runs.

Pilot Studies. Once the factors and their levels are chosen a series of pilot studies are completed to demonstrate the reasonableness of the factor levels and to determine the number of trials necessary to support the factorial experiment. The factor levels are adjusted as necessary to ensure they fall within the operational capability of both TSAR and SORGEN. The results of the second pilot study is then used to determine the statistically correct number of trials that must be run to establish an 80% confidence interval for the estimation of sorties generated. The reasonableness of factor levels is accomplished first.

Factor Reasonableness. The reasonableness of factor levels is determined qualitatively by verifying that they make sense in terms of the real system in question. This necessarily requires some degree of expertise in the airbase system. The databases used for this research were originally developed by Milt Kamins in 1987 using data collected by the RAND Corporation. They were further modified for DOD research purposes by the Air Force Center for Studies and Analysis.

The reasonableness of factors must also be verified quantitatively. This verification is conducted by making pilot runs of the databases using a simple array of treatments with the high and low factors set in different positions. The treatments used to conduct reasonability simulation runs are listed in Table 2, where lowercase letters represent the factor set to its high level, the absence of a letter represents the factor's low level, and (1) represents all factors set to their low level.

The goal of this test is to determine that the models operate properly with all factors set to their low levels, with all factors set to their high levels, with one factor set low in sequence (all others high), and one factor set high in sequence (all others low). This design should show that the factors chosen do not cause failures in the simulations. Each treatment is run for a period of 30 days and two trials with a different random seed for each treatment (see Appendix A). Upon running the reasonableness treatments, the model output is reviewed to ensure the input data did not produce unrealistic failures.

Overview of Sample Size Determination. Upon determining that the factors chosen and their established levels are reasonable, and before the fractional factorial design can be accomplished, the statistically appropriate sample size must be

Table 2

Factor Reasonableness Treatments

(1)	bcdefgh
abcdefgh	h
abcdefg	g
abcdefh	f
abcdegh	e
abcdfgh	d
abcefgh	c
abdefgh	b
acdefgh	a

calculated. The sample size in this experiment is the number of trials to conduct during each experimental treatment. To determine this number a pilot study is performed and the results are used to calculate the confidence interval for each treatment in the pilot study. The confidence interval is useful, in this case, to determine which treatment's results should be used to calculate the number of trials needed to achieve a given statistical confidence. A test of hypothesis for equal variances is also conducted; it determines whether the variance of a model under a given number of trials is statistically equal to the variance resulting from a different number of trials. The test of equal variances may be an important factor for justifying the use of a smaller number of trials when economy of simulation is a consideration.

Experimental Array for the Pilot Study. The design used for the pilot study is accomplished by making simulation runs that are 30 days in length, varying the treatments all high and all low, and varying the number of trials per run, i.e. 10, 20, and 30 trials per run. This design allows the evaluation of variance on two variables: the high and low treatments, and the number of trials. The array in Table 3 represents the experimental design used for data collection; H/10/30 is read as high treatment/10 trials/30 days, and L/10/30 is read as low treatment/10 trials/30 days. Each design point is run using a different random number seed (see Appendix A). The results of the pilot runs are then statistically analyzed.

Table 3

Experimental Array for Number of Trials Pilot Study

H/10/30	L/10/30
H/20/30	L/20/30
H/30/30	L/30/30

Statistical Analysis of the Pilot Study. Upon completion of the pilot runs from Table 3, an analysis is conducted on the results of the six treatments to determine the statistically appropriate number of trials needed to achieve an 80% confidence level in the simulations' outcome. The first analysis is the calculation of confidence intervals.

Confidence Interval Estimation. The models' reactions under different circumstances are more easily understood and observed when confidence intervals are calculated to estimate the range of values for a dependent variable. This range of values may also provide a more realistic answer to the decision-maker than a single numerical response, since certainty about model input data or their distributions is seldom absolute (25:344). The confidence interval, in this study, is an important tool in choosing the right set of treatment data to use for conducting the rest of the statistical calculations. One of the treatment differences is the number of trials conducted during each run. Varying the number of trials across treatments makes the resulting variances incomparable. The solution to this problem is the confidence interval. Since the denominator of the confidence interval calculation contains the number of trials used for that particular treatment, the variance for that treatment is normalized into a range (confidence interval) of sorties generated that can be compared to other treatments' confidence intervals.

Both SORGEN and TSAR produce a cumulative average number of sorties generated and associated standard deviation over the number of trials specified in a simulation run. These figures are used to calculate the 80% confidence interval for each treatment design point using a small sample estimation of the population mean (27:326).

$$\bar{x} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

Assumption: The relative frequency distribution of the sampled population is approximately normal. The assumption of normality is tested using a Wilk-Shapiro test. We next make a test of hypothesis for an equal population variance.

The correct set of data must be chosen from the pilot study for calculating the number of trials to use in the factorial experiment. To accomplish this the confidence intervals are graphically depicted and two decisions are made. First, the treatment group having the larger confidence intervals is chosen, and then the treatment within that group having the smallest confidence interval is chosen. This technique assures the models' variability is properly considered in two ways: 1) the treatment group with the larger variance, will return larger solutions to the number of trials calculation, avoiding the possibility of running too few trials for the given confidence level, and 2) choosing the treatment within the large confidence interval group that has the smallest confidence interval assures the number of trials solution is large enough, but not too large as to be uneconomical. The data collected from the chosen treatment are now used to calculate a preliminary number of trials for each treatment in the factorial experiment.

Calculation For Number Of Trials. After calculating confidence intervals and choosing the correct data for further calculations, the statistically appropriate number of trials is calculated. This calculation serves to justify the trial length used in making experimental production runs. The formula used for this calculation is (27:320):

$$n = \frac{4(z_{\alpha/2})^2\sigma^2}{W^2}$$

where:

W=100, the width of the confidence interval estimated with 80% confidence.

The use of 80% confidence is based on the relatively large number of experimental factors. It is expected this combination of experimental factors will produce results with relatively large variances. The use of W=100 is arbitrary and may be changed if conditions necessitate, but is reasonable (possibly tight) when considered against the number of sorties flown over a 30-day period.

The number of trials calculation produces the statistically correct number of trials to use during the factorial experiment to achieve an 80% confidence with W=100. The possibility exists, however, that the number of trials indicated will be larger than is economically feasible. Economic infeasibility may occur in two ways: 1) the number of trials indicated may be too large and thus require too much computer-time, and 2) the number, although small, may require more computer-time than is practical. In either case the economic measure for which we're concerned is the use of computer-time. In both cases, decisions must be made concerning the number of trials that will be used and the statistical confidence they provide. The question to answer is, "What number of trials is economically and statistically reasonable?" If necessary, the number of trials formula may be used in reverse to calculate the statistical confidence and value of W, given a smaller more economical number of trials. If this situation presents itself, the test of hypothesis for equal population variances is used to determine whether the use of a smaller trial size significantly diminishes the statistical confidence that can be placed in the experiment.

Test of Hypothesis for Equal Population Variances. The test of variance provides information about whether the variance observed between treatments is statistically equal. This test may be important if a decision must be made about the economy of this study. The most probable use of this test is to determine whether the variance of 30-trial data is equal to that of 20-trial data. To test the variance an F-test is conducted. This test complements the confidence interval estimation, and demonstrates whether the variance of two populations is statistically equal. Equality of variance between the 30-trial populations and 20-trial populations may indicate that 20-trial data are usable for conducting experimental runs without a significant loss in confidence. The test of hypothesis is provided and was taken from McClave and Benson. (27:415)

H_O: $\sigma_1^2 = \sigma_2^2$

H_A: $\sigma_1^2 \neq \sigma_2^2$

Test Statistic:

$$F = \frac{\text{Larger sample variance}}{\text{Smaller sample variance}}$$

the rejection region is:

$F > F_{\alpha/2}$

where:

$F_{\alpha/2}$ is based on n_{LARGER} - 1 df and $n_{SMALLER}$ - 1 df

Assumptions:

1. Both sampled populations are normally distributed.

2. The samples are random and independent.

The assumption of normality is tested during the confidence interval calculations. The assumption of random and independent samples is met based on the random nature of the models' calculations.

Experimental Design. The statistically correct number of trials determined, the experimental design used to conduct the experimental production runs is examined. This design is taken from the National Standards tables documented in *Applied Factorial and Fractional Designs* (28:253). The rationale for selecting the chosen design and its notation are discussed below.

Orthogonal Array Of Fractional Factorial. The researchers chose tentatively to use a set of eight factors based partially on our knowledge of important logistics factors, partially on the past work of Diener (12), and partially on the assumption that comparison of the two models allows this particular set of factors to be used. The selection of eight factors means that with two levels, the size of the full factorial array would be 2^8 (256), given a two-level design. To reduce the number of treatments to a more economical level, a 1/4 fractional factorial design is used. This particular design allows the measurement of all the two-way factor interactions while reducing the number of design points from 256 to 64 (28:253). This level of resolution was chosen to support on-going parallel research on metamodels conducted by Diener (12).

The fractional factorial (1/4 replication of full factorial array) is presented in Table 4. The presence of a lowercase letter indicates the use of the high factor level. Absence of the lowercase letter indicates the use of the low factor level. With the orthogonal design established, the next step is to conduct the experimental trials, collect the pertinent data, and conduct statistical analyses on that data.

Experimental Trials and Data Collection. Each of the 64 treatments in the factorial design is run using a different random number seed (see Appendix A)

Table 4

Experimental Design

1	(1)	17	abcg	33	bdefh	49	acdefgh
2	abcfgh	18	fh	34	acdeg	50	bde
3	bcdeg	19	ade	35	cfgh	51	abfh
4	adefh	20	bcdefgh	36	ab	52	cg
5	efgh	21	abcefh	37	bdg	53	acd
6	abce	22	eg	38	acdfh	54	bdfgh
7	bcdfh	23	adfgh	39	ce	55	abeg
8	adg	24	bcd	40	abefgh	56	cefh
9	cdgh	25	abdh	41	bcefg	57	aef
10	abdf	26	cdfg	42	aeh	58	bcegh
11	beh	27	acegh	43	df	59	abcdfg
12	acefg	28	bef	44	abcdgh	60	dh
13	cdef	29	abdefg	45	bch	61	agh
14	abdegh	30	cdeh	46	afg	62	bcf
15	bfg	31	acf	47	degh	63	abcdeh
16	ach	32	bgh	48	abcdef	64	defg

and the data are collected as discussed next. Each simulated treatment provides a cumulative average number of sorties generated. This is the datum that is collected from each treatment for use in the statistical analysis. Completed data collection sets the stage for the statistical analysis of the results, detailed next.

Statistical Analysis Of Results. The statistical analysis of the simulation results concludes the experimental methodology. A paired difference test is used to evaluate whether a significant difference exists between the identical treatments run on TSAR and SORGEN. The statistical formulas used to conduct this test are (27:424):

H_O: $(\mu_1 - \mu_2) = 0$

H_A: $(\mu_1 - \mu_2) \neq 0$

The test statistic is:

$$t = \frac{\overline{x_D} - 0}{s_D / \sqrt{n_D}}$$

where:

x_D(bar) = sample mean of differences

s_D = sample standard deviation of differences

n_D = number of differences = 64

The rejection region is:

$t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$

where:

$t_{\alpha/2}$ has $(n_D - 1) = 63$ degrees of freedom (df).

Assumptions:

1. The relative frequency of the population differences is normal. The normality of the population of differences will be tested using a Wilk-Shapiro test for normality.

2. The differences are randomly selected from the population of differences. This assumption is met by using the fractional factorial design which is random (i.e., there are many different fractionals that could be derived) and the

random assignment of factors to the alphabetic identifiers used in the experimental design.

Synthesis of the Qualitative and Quantitative Comparisons

In the design of this methodology, the comparisons within the three qualitative areas and the single quantitative area are presented as independent of one another. In reality, however, they are probably not independent, as depicted by the overlapping circles in Figure 5. The degree to which these areas overlap and affect one another depends significantly on the models under study, and the attention researchers conducting the study give this aspect of the comparison. Because this feature of the comparison of models is subjective, no attempt is made to qualify the relationships that must be evaluated between the four categories. This determination is necessarily left up to individual choice and no attempt is made to specify that portion of the methodology. Thus, it is important to recognize that the four areas of comparison are not independent and an overall conclusion of equivalence must be subjective and relative to the interests of the researcher/modeler making the comparison. The findings of this research will address possible conclusions from a synthesis of the qualitative and quantitative comparisons. The synthesis is based on the scope of this comparison and reflects the interests of the research team.

Summary

The designed methodology embraces a more complete comparison of two models than attempted in previous research. The qualitative comparison is important because of the need to establish a level of confidence in AMTAF that has not been achieved earlier; it also provides the information necessary for potential users to assess the traits and characteristics of AMTAF and TSAR with respect to the environment they are intended to simulate. The comparison of input and output characteristics enables potential users to assess the suitability of the two models to his/her needs. Knowing what it takes to run the model and what outputs are produced should prove useful in determining the suitability of the models to a particular purpose.

The quantitative comparison of AMTAF and TSAR establishes the level of equivalence between the two models for a specific measure of merit. This research does not compare all of the functions of either model. The analysis is constrained to evaluation of critical logistics factors and their interactive influence on simulation results, and to those factors capable of being modeled similarly while focusing on the usefulness of the proposed comparison methodology. The methodology is now tested in the presentation of findings and analysis.

IV. Findings and Analysis

Introduction

The method of investigation used in this research divides the evaluation of TSAR and SORGEN into two separate categories: qualitative and quantitative. The findings of those analyses are documented in the following sections. The qualitative findings are presented first, followed by the quantitative. The chapter concludes with a brief overall comparative analysis of the two simulations.

Results of the Qualitative Comparison

The qualitative analysis is covered in three segments which compare TSAR and SORGEN on the basis of background and documentation, features and databases, and finally on ease of use.

Model Background. The historical background and reason for the acquisition and development of a simulation model sometimes provides insight into its capabilities. Simulation models are classified by purpose and fidelity, which provide additional clues of inherent capability. A study conducted in late 1987 and early 1988 for the AFLC Logistics Operations Center assessed the potential capability of several Air Force models for "generating objective measures of merit against USAF reliability and maintainability goals" (7:i). The AFLC study proposed ***Weapon System Analysis*** as the broad classification of simulation models that are used to predict and evaluate military operational capability. There are three specific areas of application, under this broad classification, that pertain to Air Force weapon systems: logistics, the airbase, and the mission. The researchers adapt this classification scheme to compare

TSAR and SORGEN. The qualitative analysis begins with the findings of the investigation of TSAR's background.

TSAR Simulation Background. Rand Corporation's development of the Theater Simulation of Airbase Resources (TSAR) began in the late 1970s under the Project Air Force Resource Management Program project entitled *Strategies to Improve Sortie Production in a Dynamic Wartime Environment* (14:iii). TSAR "is a Monte-Carlo discrete event-driven simulation model that analyzes the interrelations among available resources and the capability of the airbases to generate aircraft sorties in a dynamic, rapidly evolving wartime environment" (18:1). The simulation evolved through the 1980s as RAND enhanced the basic model for the Air Force to include more realism, e.g., rapid runway repair, alternate equipment repair procedures, and increased task times for personnel using chemical protective equipment (13). The most recent version of the simulation, which is used in this research, is an update to the 1985 version. The purpose of the simulation in its current form is capsulized in the following quotation:

> The TSAR (Theater Simulation of Airbase Resources) model simulates a system of interdependent theater airbases, supported by shipments from the Continental United States (CONUS) and by intratheater transportation, communication, and resource management systems. By capturing the interdependencies among 11 classes of resources, the simulation permits decision-makers to examine the implications of many possible improvements in terms of their effects upon the sortie generation capabilities of the system of airbases. The simulation also allows examination of the effects of damage inflicted by enemy airbase attacks using both conventional and chemical weapons and the results of efforts to restore operations. (18:1)

TSAR is a high fidelity simulation with broad capabilities to simulate many of the operational aspects of a theater system of airbases. Operational impacts are represented via user-specified attrition and battle damage rates enabling simulation of lost and damaged aircraft and the resulting impact on the logistics support infrastructure. The AFLC report classifies TSAR as a sortie generation model under the airbase application (7:2-4). However, TSAR possesses capabilities which permit the extensive evaluation of logistics capabilities associated with the operation of a system of airbases, embracing large portions of the logistics application. A more encompassing single classification is needed to appropriately classify the TSAR model. Therefore the classification of *Airbase Operability Model* is proposed which adds theater and CONUS logistics and logistics policy simulation. There appears to be no concise system of classification that is widely used and accepted for military models due in part to the fact that many models are expanded over time to include aspects of the environment that have a direct impact on the specific function originally simulated. This causes the model's classification to shift and contributes to a general misunderstanding of the relationships of the various models with similar capabilities. Next SORGEN's background is investigated.

SORGEN Simulation Background. SORGEN is the sortie generation module of the All Mobile Tactical Air Force (AMTAF) suite of models. The AMTAF user's guide states:

> The All Mobile Tactical Air Force (AMTAF) suite of models was developed as a balanced weapon systems evaluation model to enable the analyst to measure and prioritize the many activities associated with deploying and supporting weapon systems. The AMTAF system is partitioned into four models which simulate airbase operations, airbase attack, logistics, and sortie effectiveness. The modular nature of the system allows the user to focus on one area at a time while the model linkages enable end-to-end analysis from supportability and operability to mission effectiveness. (2:1-1)

The AMTAF suite of models was developed for Aeronautical Systems Center Mission Area Planning office (ASC/XRS) to support effectiveness analysis for Air Force missions. This analysis is used in developing requirements and weapon system acquisition planning. The AMTAF effort began in the mid-1980s. The completed model was delivered in 1990 (26:20).

Ball Systems Engineering Division, AMTAF's developer, classifies SORGEN as a sortie generation model which focuses on the simulation of airbase operations. Its "objective is to model the sequential flow of events and the resources required to generate sorties" (2:5-1). SORGEN was designed as an intermediate fidelity event-driven Monte Carlo simulation. SORGEN simulates a main operating base supporting one or more dispersed operating bases.

Referring again to the classification scheme proposed by the AFLC study, SORGEN falls in the sortie generation area of the airbase application. SORGEN simulates base repair, both on and off equipment, but does not simulate CONUS resupply, centralized intermediate repair (CIRF), depot repair, and transportation of resources. These features are simulated in the logistics simulation (LOGSIM) module of AMTAF. SORGEN alone does not fit the classification of an airbase operability model proposed earlier. Concluding that TSAR and SORGEN differ slightly in age and classification, attention turns to the comparison of the respective sets of documentation. Table 5 presents a model characteristics summary.

Documentation Comparison. Model documentation is critical to the successful post-development use of a simulation model. The user and programmer/analyst manuals are the only perpetual source of information pertaining to the purpose and use

Table 5

Model/Module Characteristics

Model/Module	Fidelity	Documented Classification	Study Classification	Development Began
TSAR	High	Sortie Generation Model	Airbase Operability Model	Late 1970s
SORGEN Module	Intermediate	Sortie Generation Model	Sortie Generation Model	Mid 1980s

of the simulation model. The following sections document observations and experiences using the documentation in this research.

TSAR Documentation Analysis. The documentation for TSAR and TSARINA is contained in a four volume set of manuals that describe the two simulation models. The first three volumes are dedicated to TSAR and the fourth to TSARINA which also supplements the AMTAF documentation.

Volume I "of the *User's Manual* provides a full description of the logic used in the TSAR model, as well as an understanding of the interrelations among the many elements of the logic" (18:iii). The documentation is structured to serve four classes of readers: those seeking an overview of the simulation's capabilities, those with a knowledge of programming seeking a full understanding of TSAR's logic, those preparing input databases for simulation runs, and finally those interested in modifying or correcting the existing program logic (18:12).

Volume II of the *TSAR User's Manual* deals primarily with the details of the input database. Included with these explanations are example data entries, each with an explanation, which enable a thorough comprehension of the database. Volume II was the most used of all the TSAR documentation during this research.

Volume III contains the programmer/analyst information and was the least used of the TSAR references. However, the glossary of terms and the time uncertainty distribution definitions were referenced as the database was interpreted. The researchers skills are inadequate to assess the adequacy of programmer/analyst information.

Volume IV covers TSARINA, the airbase attack feature. Since the attack feature is not used in this research the manual is not evaluated.

There is no distinct division in the TSAR documentation between user and programmer/analyst sections. While the author's intent to serve four types of readers is achieved it is often not clear where one level of detail ends and the next begins. During the conduct of the research considerable time was spent searching for the details and information related to database content, running the model, and debugging the database. The manuals each contain a table of contents but no index. There were numerous occasions when a comprehensive index would have been useful. The high degree of fidelity in the TSAR model is well documented.

The structure of the documentation follows a logical format of general to specific that permits the reader to delve into whatever level of detail is required. For the programmer/analyst this could require reading the entire manual set before locating the needed information. Again, an index to complement the table of contents would be useful. The investigators found the level of detail more than adequate for the purposes of the research. The diagrams, tables, and illustrations improve understanding. The example data and sample problem are useful, providing useful details which otherwise would be missed or misunderstood. The documentation is clear and understandable. The complexity of the model is clearly illustrated and explained thoroughly, however, more thorough coverage of debugging databases is needed. Next the AMTAF documentation is analyzed.

AMTAF Documentation Analysis. The documentation for the sortie generation (SORGEN) module of AMTAF is contained in a set of four manuals that cover the entire suite of AMTAF models, i.e., SORGEN, the Logistics Simulation (LOGSIM) model, the Mission Area Simulation to Evaluate Requirements (MASTER) model, and the TSAR INputs using AIDA (TSARINA) model. AMTAF documentation for TSARINA is supplemented by manuals from RAND Corporation, the TSARINA developer.

The *AMTAF System User's Manual* provides an introduction and overview of the AMTAF suite of models. It states that it "provides a description of model inputs and documents the program's features and operation. The manual is intended for use by personnel who are not programmers and should provide sufficient guidance for model operation" (2:1-2). The manual begins with an brief background, historical perspective and overview of the AMTAF suite of models. Section 5 of the *User's Manual* documents the SORGEN module with an overview, a brief explanation of the program structure, an explanation of the major simulated phases of the sortie generation process, and the resources used to support it. Two example cases depict alternative uses of SORGEN. The simulation software obtained for this research came with these example databases and aided understanding the simulation's features. The appendices to the manual document the example databases and provide examples of output reports for SORGEN, LOGSIM, and MASTER. The input databases for the AMTAF suite of models are managed by VICS, the VERAC Information Control System, a relational database manager. Instructions for using VICS is contained in Appendix A of the manual.

The *AMTAF System Programmer/Analyst's Manual* was used very little during this research. The manual and its accompanying volume of appendices provide the program structure and logic. Again the researchers' lack of ability as programmer/analysts negate evaluation of the technical quality or usefulness of these volumes. The SORGEN section of the appendices includes the database schema or structure files used by VICS and are useful in identifying and correlating data input requirements between SORGEN and TSAR. The VICS data element labeling feature proved useful in the creation of the database mapping mechanisms to be discussed later.

The overall structure of the documentation follows a logical format of general to specific. The partitioning of the information into user and programmer/analyst categories is logical and appropriate. Portions of the programmer/analyst's documentation are useful, however, most users would probably not use the information. The investigators found the level of detail in the SORGEN documentation to be marginal. A more thorough explanation of the airbase functional relationships is needed as is a section covering database debugging. The diagrams and figures presented are appropriate to the present level of detail. The example databases provide a useful tool with which to gain basic knowledge and experience. Overall the documentation is clearly written and easily understood given the limitations noted.

Comparison of TSAR and AMTAF Documentation. The documentation for the TSAR model is more detailed, attributable in part to the higher degree of fidelity in the model. The AMTAF documentation is more clearly partitioned into user and programmer/analyst segments. Neither set of documentation is indexed, which increases the effort required to search out information. Both sets of manuals effectively use illustrations, tables, and figures. The TSAR and AMTAF manuals are organized logically and the level of detail progresses from general to specific. Both are readable and easy to comprehend. The greatest difference noted between the sets of documentation is in the level of detail. The TSAR manuals present more detail in the sections believed to be written to the user, whereas, AMTAF user's sections are more general in nature. An overall perception is that the TSAR documentation is more detailed and therefore better supports post-development employment. Neither models' manuals cover database debugging adequately. Ball Systems Engineering Division personnel assisted the researchers, on several occasions, in achieving the level of knowledge needed to build and debug the SORGEN databases when the documentation alone was insufficient. Having gained the requisite level of model familiarity from analyzing the documentation the qualitative analysis next compares the features and databases.

Features Catalog and Database Comparison. A comparison of features between TSAR and SORGEN provides a ready reference from which to base an assessment of the models' suitability for a particular use. The comparison of the databases provides insight into the data required to drive the simulation models. Comparisons of model features and databases are not sufficient alone to select a simulation model, but they do expand the understanding of the models and their capabilities.

TSAR/SORGEN Features Table. During the course of this research, 128 specific model features associated with airbase operability and sortie generation were identified. The model used to derive the baseline list of features is TSAR, since overall it has higher fidelity than SORGEN. In the AMTAF suite of models, the features are split between LOGSIM, the logistics simulation and SORGEN. Since this research compares only SORGEN and TSAR, no attempt is made to comprehensively catalog LOGSIM's features; however, some of its features are referred to in the SORGEN documentation and are therefore included in the features table. The features table is located in Appendix B. Even though TSAR and SORGEN share many common features, significant differences are also present.

Significant Differences Between TSAR and SORGEN. While there are numerous differences between the two models, many are simply differences in fidelity where more or less input data are required or where there is a greater level of detail in the output reporting. There are features, however, representative of actual airbase operation, which the researchers believe have considerable influence on the simulations' prediction of sorties generated. These differences are discussed in Appendix C. While there are numerous differences between the two simulation models, 71 of the features are common. The simulations both emulate a system of theater airbases operating one or more types of aircraft with maintenance and supply functions necessary to the support of combat operations. Consumption and resupply of munitions, fuel, parts and support equipment are simulated along with attack and attack recovery. Finding commonality in features led the researchers to expect similar findings in the input database content, which is the next area of investigation.

Database Correlation and Mapping. TSAR and SORGEN exhibit distinct differences in their respective input database formats. TSAR uses an 80-column card format while SORGEN uses a relational database format. The process of correlating the TSAR and SORGEN databases required the development of a data mapping mechanism that matches the TSAR card type and column number to the corresponding SORGEN database and database relation.

TSAR categorizes data by card type. There are 117 card types which are grouped into 15 categories. The categories partition the data functionally, e.g., card types 17/1 through 19 contain data that describes the airbase, card types 29, 29/88 and 30 contain aircraft maintenance scheduling data. Card type 40 contains the data related to attack, a feature not used in this research, and is not included in the cross-indexing of the databases.

SORGEN uses 10 databases to organize the input data categorizing it functionally into distinct groups. Individual relations within each database further partition the data into elements. For example, the BASE database contains the data that describes the airbase and its aggregate resources, such as runways, taxiways, shelters, parts, aerospace ground equipment, personnel, etc. Each of these categories are relations. The naming convention used in VICS, the relational database manager, relates data by database and relation, e.g., BASE/PERSONNEL is the BASE database

PERSONNEL relation. Within the relations there are multiple records which define the individual types and quantities of resources. The ATTACK and TSRNA_EQUIV databases are unique to the attack feature of the simulation which is not exercised in this research and are therefore not included in the database mapping.

The SORGEN database structures are drawn into a columnar format using the VICS database manager, relation, and data labels. The corresponding TSAR card type and card column number are presented in parallel with explanatory comments as needed. Appendix D contains the database map that correlates the SORGEN database to the TSAR database in SORGEN database/relation order. Appendix E contains the database map that correlates the TSAR database to the SORGEN database/relation in TSAR card type sequence.

The investigators use these mapping mechanisms extensively to create the equivalent databases for the quantitative comparison of the models. A significant portion of the databases translate directly. The use of common data in multiple locations and the aggregation and segmenting of the data between the databases negates easily calculating the percentage of commonality between the two. Having correlated and created equivalent input databases, the research then focused on the actual use of the models.

Useability Assessment. The ease with which a simulation model can be learned and used contributes directly to the extent to which it will be used. In the following analysis the researchers present their experiences using TSAR and SORGEN. As was stated earlier, neither researcher had prior experience with either of the simulations, which minimizes potential bias of the subjective comparison. The comparison considers six areas believed critical to the useability of a simulation model: 1) ease of learning, 2) ease of database development, 3) availability of database debugging tools, 4) ease of implementing an experimental design, 5) computer run time, and 6) adequacy of output data products. A discussion of the individual models is presented followed by a brief comparison of the two simulations.

Ease of Learning. The process of learning the models began at the outset of this research. The primary tool available to assist in this learning process is the documentation from each of the two models. In the case of TSAR, the documentation was supplemented with actual experience from the Diener research (12,13). The researcher assisted in getting the model mounted on the host computer and setting up the database and output report directories. He also provided instruction in the operation of the simulation, helped debug the database, and assisted in the initial interpretation of the output data.

In the case of SORGEN, the model documentation was supplemented by technical assistance from Ball Systems Engineering Division (BSED). They provided a training session where AMTAF was demonstrated, input databases described, module interfaces explained, and output reports illustrated. In addition to the training, BSED provided technical assistance and consultation throughout the course of the research.

Learning the models proved to be an iterative process, continuing throughout the research. The exact point where the team became proficient operating the models is unclear, but likely occurred during the final stages of the research when the models were being used to run the pilot and experimental trials. The early research efforts focused on the database development versus the actual operation of the simulations. Given the resources made available to this research team, proficient operation of the models can be achieved in a matter of a few days, assuming no database development is required. Gaining a thorough understanding of the respective databases is not a trivial undertaking; it takes considerable time and effort and is believed necessary to properly employ the models. The VICS database management feature of the AMTAF suite made understanding the SORGEN databases easier. The on-line data description feature provides a concise description of the required data and data format. The 80-column card format of the TSAR database is more difficult to interpret. However, once familiar with the various card types and content, use and interpretation of the database is not overwhelming.

Overall, SORGEN has the edge in ease of learning, while neither simulation is difficult to operate, assuming the prospective user is proficient in the UNIX and Digital Control Language (DCL) operating systems. However, neither qualifies as "user-friendly" by today's software standards.

Ease of Database Development. Developing equivalent databases for the two simulations proved to be the most time-consuming aspect of the research. While not difficult, the very large baseline database made the task complex. The majority of the effort focused on cataloging the features employed by the TSAR F-15 database and then translating those features into SORGEN. The database maps developed earlier proved critical to this effort. Starting with an existing database negates comparison of database development activities, however, the large amount of data contained in the baseline database represents a considerable data collection effort.

Inputting the translated data to the SORGEN databases proved time consuming even given the utility of VICS. One large block of data related to the aircraft maintenance task requirements was translated and input using a conversion utility developed by BSED from a specification developed by the researchers. One observation became readily apparent as the SORGEN database evolved: database development necessarily requires a thorough understanding, not only of the model(s) but also the structure of the databases, units of measure, and program functions controlled by the input database. Once developed the databases were individually debugged.

Availability of Debugging Tools. During the debugging effort errors in the databases were discovered and subsequently corrected. Exercising the documented error detection features on both models proved inadequate for locating all of the errors that prevented the simulations from running. While extended error detection capability exists in both simulations, neither is sufficiently documented for use by this research team. Debugging TSAR proved to be relatively easy since the researchers had a known operable database to use as a test platform. Using the original database, segments of the changed database were substituted sequentially until all had been tested. Assembling the new segments produced an operable research database.

Debugging SORGEN proved more difficult. While relying on the input check and debug capability to analyze the input database, the research team encountered conditions which escaped detection and caused the simulation to terminate prematurely. BSED personnel loaded the databases on their computers and used software development and debugging tools, which were not part of AMTAF, to locate the errors.

Given the different avenues used to debug the databases a comparing the models' debugging tools was not accomplished in the context of this research. Had both models been required to run newly developed databases that were debugged by the research team without outside assistance then a meaningful comparison might have been possible. Completing the database debugging activities set the stage for the next phase of the research; incorporation of the experimental design.

Ease of Implementing an Experimental Design. This comparison provides an insight into using the models for "what-if" analysis where several options are to be considered. The experimental design of eight factors at two levels required building several versions of the input databases to run the various combinations of factors in the treatments specified in the fractional factorial array. Structuring the TSAR database required segmenting it by the card types that contained the individual factors. High and low versions of each factor are then constructed and stored. The initial portion of the TSAR database contains the simulation initialization and set-up information and further specifies which database segments to use. Experimental treatments, made up of specific combinations of the experimental factors, are attained by specifying the appropriate version of each database segment.

Formulating the experimental databases for SORGEN proved more difficult since the individual factors are contained in individual database relations. The BASE database for example contains four factors, which at two levels each produces 16 possible combinations. This required 16 versions of the BASE database to be built.

The individual versions were then modified to set the various relations to their required levels. This same convention was used to build the BASE_MODS database. The remaining two factors, mission assignment and filler aircraft, were set in the SCENARIO database. This database also identifies the version of the BASE and BASE_MODS databases to be used for a particular "scenario." The random number seed, number of trials, and output report options are specified in the CONTROL database. Eighty-eight unique CONTROL and SCENARIO databases were ultimately created to accomplish the experimental, reasonableness, and variability runs. The relatively short time required to create the various database versions demonstrated the utility of VICS as an input database management system.

Even though TSAR and SORGEN exhibit distinct differences associated with doing experimental or "what-if" analysis, neither was extraordinarily difficult to use. VICS makes manipulating the SORGEN databases easier than manipulating the TSAR databases with a computer text or line editor. Overall the models were found to be approximately equivalent with respect to incorporating the experimental design. Run time analysis is the next portion of the qualitative evaluation.

Run Time Analysis. One of the comparison factors planned for this research was simulation run time. To directly compare the two simulations, they must be operated on the same computer system. Circumstances forced us to operate the models on different computers under different operating systems which negates direct comparison of run times. However, this measure of performance is believed useful in assessing the individual simulations' useability.

The TSAR simulation runs were made on the AFIT ELXSI computer. TSAR executes several runs simultaneously which permits multiple submissions during a single computer session. Setting up runs in a queue also helps reduce the amount of time needed to execute the runs. The TSAR runs averaged 168.7 minutes each over 61 experimental trials. The run time data for the remaining three trials were lost to an archiving failure. The average time is computed from CPU times recorded in the normal output report. Run times ranged from 102.33 minutes to 300.57 minutes with a standard deviation of 44.08 minutes.

The SORGEN simulation runs were made on the AFIT VAX computer cluster. SORGEN runs only a single simulation at a time which consumes a significant amount of computer time when multiple runs are required. However, using batched submissions the 64 experimental trials were completed in one 56 hour period extending over three days. The simulation runs averaged 33.86 minutes each over the 64 experimental trials, computed from CPU times recorded in the output report. The times ranged from a high of 47.58 minutes to a low of 24.59 minutes with a standard deviation 6.5 minutes.

While TSAR took much longer to run the trials, the fact multiple trials could be run in parallel roughly offsets the shorter run times for SORGEN which ran the trials serially. Again, the reader is cautioned that a direct comparison is impossible due to the differing computers and operating systems. Assessing the output reports is the next topic of discussion as the qualitative analysis continues.

Adequacy of Output. Both simulations provide a high degree of flexibility in output report selection. The minimum output was used in each instance to conserve computer storage and to minimize the consumption of paper should printing some or all of the reports be necessary.

TSAR has preformatted output reports that are selected via the input database. The numerous options and levels of detail provide the prospective user a seemingly endless array of alternatives. TSAR has post-processing capability in addition to the pre-defined reports which permits analyzing all data that TSAR writes to disk and offers options that are believed to suit even the most advanced user's needs. Summary statistics are provided on the numerous events contained in the formatted reports as are daily and cumulative statistics for sorties flown. In addition to the preformatted "Normal" TSAR output report, which contains textual data labels and column headers, two additional unlabeled data-only reports are written for each simulation run. The "long" and "short" reports contain data with no text labels and can be read directly into analytical tools such as spreadsheets and SAS routines.

SORGEN uses preformatted reports exclusively, permitting the user to specify the quantity of output desired via the CONTROL database CONTROL relation. The available options provide the user a rich selection of report combinations which should meet the majority of most user's needs. The output reports provide summary statistics on the multitude of reported events. Daiiy and cumulative statistics are provided for sorties flown, both total and by mission. Advanced modelers having need for data not contained in the preformatted reports face some difficulty since SORGEN has no documented post-processing capability. A second shortcoming is that the SORGEN output reports are not in a format that is directly usable in other analytical tools such as spreadsheets or SAS routines. Post-trial analysis of simulation results requires manual extraction from the preformatted reports for input to other programs.

The report options for the two simulations are too numerous to list as part of the findings. They are documented in the individual simulations' manuals. The reader should refer to the appropriate manuals for a comprehensive discussion of the available options. It is the opinion of this research team that the available output reports are adequate to the needs of most model users, with the exceptions noted above. The following summary concludes the model useability assessment.

Useability Summary. The useability of TSAR and SORGEN are similar. While each has specific weaknesses, each also has strengths. The VICS database manager makes the use of SORGEN easier with respect to database manipulation. SORGEN lacks a post-processing capability and does not produce data files that can be used directly in other analytical tools. Manipulation of the TSAR database is more cumbersome than SORGEN, however, TSAR's post-processing capability enables customizing output reports for specific purposes. Each model exhibits roughly the same level of useability, neither being overly difficult to employ. Running the simulations on different systems negates comparing the models' runtime, however, this team did not find run time to be a limiting factor in the completion of the necessary runs to support the quantitative analysis. Preparation and manipulation of the respective databases, while different in process, proved to be workable. The models are judged to be approximately equal overall with respect to useability. Table 6 provides a summary of the useability comparison. The qualitative assessment concludes with a summary of the observations and findings.

Table 6

Useability Comparison

Useability Category	TSAR Model	SORGEN Model
Ease of Learning	Considerable effort due to model complexity and database format	Slightly easier due to less complexity and VICS database manager
Ease of Database Development	Relatively easy. Available database used and modified for equilibration. New database development would be significant effort.	More difficult due to translation process. Process estimated as approaching new database development complexity.
Availability of Debugging Tools	Available but too poorly documented to be used by most users. Used segment substitution as work around.	Available but too poorly documented to be used by most users. Resorted to outside assistance to complete.
Ease of Implementing Experimental Design	Relatively easy as database is logically segmented by functions. Permits substitution of segments with different settings.	Large task due to number of versions required to implement design. VICS utility demonstrated by short time required to complete.
Run Time Analysis	Long run times, model permits multiple runs to be made in parallel	Shorter run times, model permits batched submissions but runs only one run at a time.
Adequacy of Output	Preformatted, data-only, and post processing capability. User controlled. Should meet all users requirements.	Preformatted only. User controlled. Should meet majority of all users requirements.

Qualitative Summary. Thus far the research has compared TSAR and SORGEN qualitatively. The background and purpose of each has been investigated. The two models have been placed into a classification scheme where they were found to be slightly dissimilar. TSAR possesses logistics features not found in SORGEN. The documentation comparison revealed shortcomings in each set of manuals. TSAR's documentation, while more comprehensive overall, lacked the clear partitioning noted in the AMTAF manuals. Both sets of manuals lack indexes and database debugging instructions.

A comparison of 128 simulation features revealed that TSAR and SORGEN share 72 of them. There are ten notable differences in capability that are believed to have a significant influence on model performance. The database comparison revealed differences in format, yet commonality in content. The database mapping mechanisms produced as a by-product of the database comparison were subsequently used to derive the equivalent experimental databases.

TSAR and SORGEN were assessed as approximately equal in useability when comparing factors dealing with learning to run the simulations and the complexities of developing and manipulating the input databases. The models' run time, while measured, cannot be compared since the trials were run on different computers. The comparison of output data products concluded that TSAR's post-processing capability and data-only reports provided capability beyond that present in SORGEN.

Clearly there are differences between TSAR and SORGEN. Both models have strengths and weaknesses. It is the opinion of this research team that SORGEN is slightly easier to learn and use while TSAR is of higher fidelity and better documented. Neither has adequate debugging tools. Table 7 provides a summary of the qualitative comparison. Next, the results of the quantitative comparison of TSAR and SORGEN.

Table 7

Qualitative Comparison

Comparison Category	TSAR Model	SORGEN Model
Background	Developed by RAND circa 1980 - 90	Developed by Ball Systems 1985 - 90
Purpose	Sortie generation and airbase operational analysis	Sortie generation within a suite of models used for overall weapon system analysis
Classification	Airbase Operability	Sortie Generation
Fidelity & Level of Performance	High, theater system of airbases	Medium, theater system of airbases
Documentation	Comprehensive, lacks clear partitioning, debugging instructions and indexing.	Clearly partitioned, lacks detail, debugging instructions, and indexing.
Features	128 specific	76 specific (72 common to TSAR)
Input Database	80 column IBM card format, 117 card types 15 categories	Relational database, 10 types with multiple relations
Useability	Workable, learning model and database difficult due to complexity. Preformatted, data-only, and post processing output.	Workable, learning model and database eased by VICS database manager. Preformatted output only.

Quantitative Comparison

Comparison of TSAR and SORGEN on a quantitative basis consists of several phases: 1) establishing a measure of merit common to both models, 2) developing databases that are equivalent to drive each model, 3) developing an experimental design that extensively exercises the models, 4) making pilot runs with each model to ensure the chosen factors and levels do not cause fatal errors, and 5) running the experimental trials and analyzing the results. The final results of each model are compared statistically to determine the degree of equivalence that exists between the two simulations. The following sections document the quantitative research and tabulate the results.

Measure of Merit. As discussed in Chapter III the quantitative comparison of two simulation models requires, ideally, common numerical measures of overall performance. While alternative means of evaluating performance exist, such as multivariate comparisons, a simpler analysis was chosen for this research. The research team believes that the effect of all aspects of the airbase simulation are observed in a single measure: sorties flown. The *AMTAF User's Manual* says of SORGEN: "The primary measure of merit is *sorties* generated" (2:5-1). The *TSAR User's Manual* says that: "TSAR is a Monte Carlo discrete-event simulation model that analyzes the interrelations among resources and the capability of the airbases to generate *aircraft sorties* in a dynamic, rapidly evolving wartime environment" (18:1). The findings of the qualitative comparison of TSAR and SORGEN conclude that each is intended to simulate the operation of a military airbase, the end product of which is sorties generated, i.e. the number of aircraft flown. Restating the assumption that the simulation models are structured such that the interactive airbase functions or factors influence overall sortie production, and since each simulation produces a quantitative estimate of sorties produced, cumulative sorties generated over a 30-day period is the chosen measure of merit for quantitatively comparing TSAR and SORGEN.

The preliminary factors chosen during the development of the experimental methodology proved common to each model in the qualitative comparison and are carried forward into the database development effort. The eight experimental factors are:

1. Aircraft
2. POL (fuel)
3. Munitions
4. Missions
5. Personnel
6. Spares (parts)
7. AIS (avionics intermediate shops)
8. Support equipment

Figure 11 depicts the relationship of the chosen factors to the measure of merit and are shown in the shaded ellipses. All of the ellipses in the figure represent simulation features present in one or both of the models. The modeling environment encompasses all of Figure 11. The researchers have illustrated a boundary for the airbase only to clarify the conceptual airbase environment. The boundary and environment designations are not intended to illustrate the entire simulated environment of either model as each has capabilities which extend beyond the limits of this construct. Having chosen the measure of merit and established its relationship to the chosen factors, attention turns to the development of the databases needed to run TSAR and SORGEN.

Database Translation and Equilibration. Using the database maps, developed as part of the qualitative analysis, equivalent databases are developed for TSAR and SORGEN. Dissimilarities in the simulation models and their databases are handled consistently. When a feature is encountered that is not in both models or is not obviously replicated equally, the lesser capability becomes the standard and the model possessing the extended capability in that particular function is constrained to a level that, as nearly as possible, equates it with the lesser capability. Where differences in level of detail are encountered, the data inputs are aggregated or

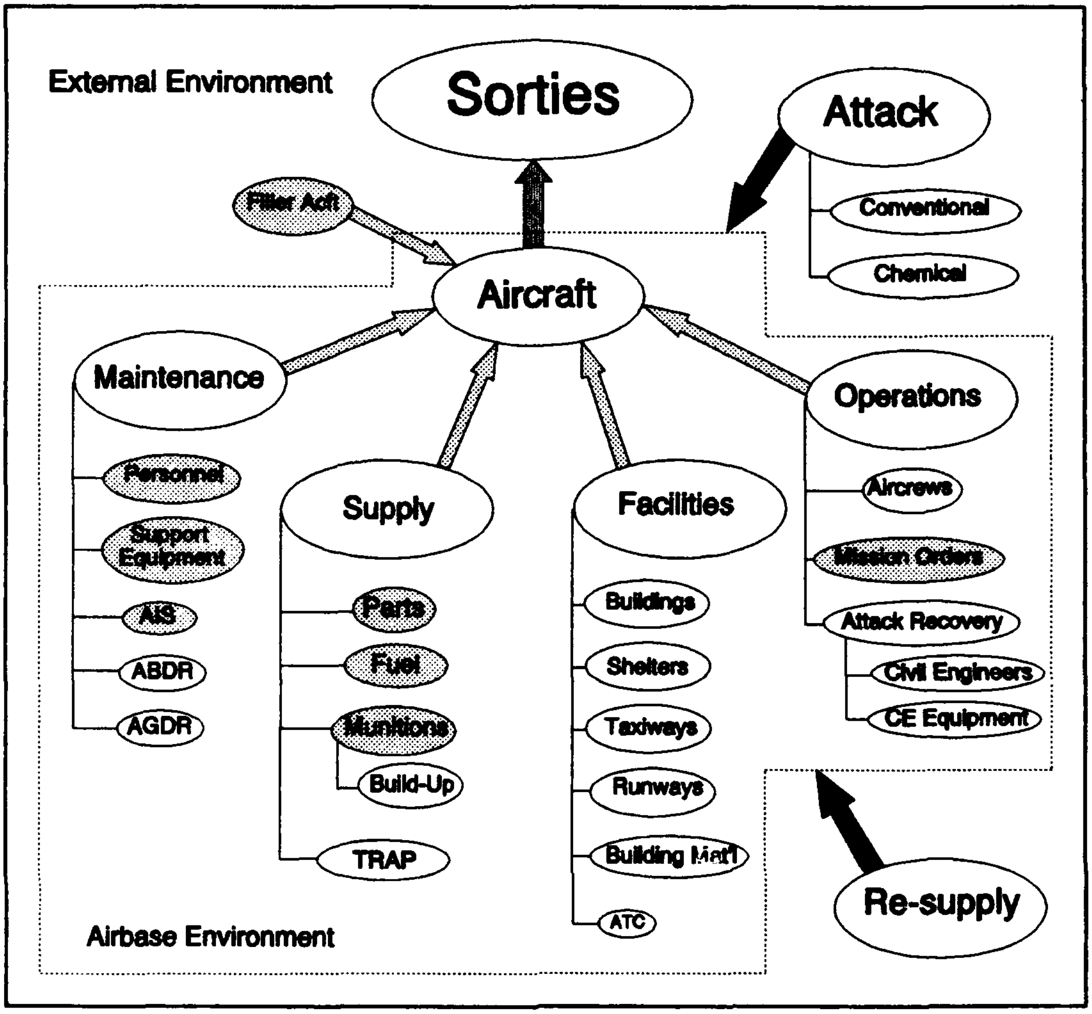

Figure 11. The Relationship of Experimental Factors to Sorties

disaggregated, to exercise the maximum number of common functions. These steps are necessary to exercise TSAR and SORGEN as thoroughly as possible within the context of this research. A total of 31 changes had to be made to the TSAR F-15 database to enable formulation of equal SORGEN inputs. These changes are limited to a no attack case and are documented in Appendix F.

The databases are structured to represent a single airbase, initially equipped with 72 aircraft. The simulation period is 30 days with no resupply of parts, personnel, or support equipment. Periodic resupply is provided for fuel and munitions at scheduled intervals and varied between high and low experimental treatments. The base incurs no enemy attack. The aircraft incur a decreasing operational attrition rate, initially 1.2%, declining over two days to 1.0%, where it is then held constant for the duration of the simulation. The aircraft incur no battle damage and the corresponding maintenance tasks are disabled. The base is tasked with five air-to-air type missions with varying flight durations, flight sizes, priorities, missile (munition) loads, and preparation times. Munition expenditure is assumed to be 100%. The sortie demand rate is varied between high and low experimental treatments.

The resultant databases are archived in AFIT/LAL. Researchers desiring to expand on this work or wishing to replicate this research should contact Lt Col David A. Diener at (513) 255-5023, or the researchers at their permanent addresses documented in the vitas. The reader is reminded of the database maps in the appendices that provide cross-tabulation indexes for the databases. With equivalent databases complete, the next step is to select the experimental factors and set their levels.

Experimental Factor Selection. As discussed in Chapter III several factors and factor levels representative of processes of an airbase operation are used to build the experimental design for this research. The factors are clearly defined in the input databases and possess values that are varied to facilitate experimentation.

While it was our intention to individually test all of the factors at each level in each model for reasonableness, and to adjust them as needed prior to the conduct of the experiment, it was not possible to test them in SORGEN prior to accomplishing the TSAR experimental trials. Two conditions drove the need to alter the approach. First, SORGEN in the original configuration had memory limits too small to accommodate the input database developed for the research, a condition which could not be corrected early enough to proceed as planned. Second, the inability to satisfactorily run TSAR on the AFIT SUN computers required the experimental runs to be made on the AFIT ELXSI computer system prior to its deactivation. The team therefore had to assume that since the chosen experimental levels ran reliably in TSAR, they were suitable for the overall comparison. Further, since TSAR is the baseline simulation against which SORGEN is tested, it is assumed that factors shown useable on TSAR must also be useable on SORGEN for equivalence to be assessed.

Pilot Studies. Two pilot studies are conducted as a part of this research. The first tests the reasonableness of the factors chosen for the experiment. The second provides data needed to determine the statistically correct number of trials to run during the factorial experiment in order to achieve an 80% confidence in the results. The results of both pilot studies follow; the results of the reasonableness study are presented first.

Factor Reasonableness. Prior to beginning the actual experimental runs, the equilibrated databases must be tested to assure they are reasonable. Reasonable in this sense means that the databases will work with the models and do not produce errors that prevent the trials from being completed. Table 8 contains the experimental treatments used for the reasonableness tests and the resultant sorties generated, these tests are based on two trials. Each of the reasonableness treatments is tested using a different random number seed. In TSAR each of the reasonableness treatments operated completely, without terminating early, and produced complete end-of-run reports. The same reasonableness design is used to test the experimental factors in SORGEN. The results of running the reasonableness treatments in SORGEN produced runs that did not terminate early and produced complete end of trial reports.

Table 8

Reasonableness Array Design

Treatment	TSAR Sorties Generated	SORGEN Sorties Generated
(1)	3151.0	1690.6
abcdefgh	3023.5	3101.2
abcdefg	4183.0	3011.6
abcdefh	3176.5	2740.4
abcdegh	3699.0	2115.6
abcdfgh	4437.5	2731.0
abcefgh	3377.0	2718.4
abdefgh	3433.0	2843.8
acdefgh	2741.0	2853.4
bcdefgh	2477.0	2718.4
h	2984.0	1827.8
g	3489.5	1917.4
f	3523.0	2369.8
e	2876.0	1908.8
d	3064.0	1762.8
c	2897.5	1930.2
b	3080.0	1827.2
a	3192.5	1831.6

Every effort was made prior to conducting the reasonableness treatments to ensure the databases used for TSAR and SORGEN were equivalent and properly debugged. The success of the reasonableness treatments now allows us to assume the databases are equivalent, are properly debugged, and will function properly in both TSAR and SORGEN across the combination of factors tested in the experimental trials. The next step in conducting the quantitative analysis of the two models is to determine the statistically appropriate number of trials to include in each experimental treatment.

Sample Size Determination. The second pilot study conducted is used to determine the statistically appropriate number of trials needed to achieve an 80% confidence in the models' estimate of sorties generated. Data are collected from the pilot study and used to calculate the confidence intervals for each of the treatments run in the pilot study. A decision is then made about the appropriate data to use for the rest of the statistical calculations and a preliminary number of trials is calculated based on the chosen data. The preliminary number of trials indicates that further calculations must be made in order to determine the number of trials to use in the factorial experiment. The preliminary number is not economically feasible. A test of hypothesis is conducted to determine if the variances observed between 30-trial and 20-trial data are equal. Based on these findings a final determination is made on the number of trials to use in the factorial experiment.

Experimental Treatments for the Pilot Study. The treatments used to determine the number of trials to use during the factorial experiment are found in Tables 9 and 10. The results of each treatment, mean sorties generated and standard deviation are also provided. The treatment used in each of the three high treatments was "abcdefgh," (all factors at the high level) and the treatment used in each of the low treatments was "(1)," (all factors at the low level). The convention H/10/30, L/10/30, etc., is read as an all high treatment, 10 trials, 30 days, or an all low treatment, 10 trials, 30 days. The first statistical study conducted is confidence interval estimation.

Confidence Interval Estimation. The confidence interval is calculated for each of the six treatments in both TSAR and SORGEN, the results are presented in Table 9 for TSAR and Table 10 for SORGEN. The expected response was observed; as the number of trials is increased the confidence intervals become narrower. The variability present between the treatments is more easily observed if the confidence intervals are displayed graphically. This is due to the fact that different trial sizes are used which renders the resulting standard deviations incomparable. The confidence interval normalizes the variance of each treatment into a range of sorties generated and allows a comparison of treatments. The confidence intervals for each of the variability treatments is presented in Figure 12.

The confidence interval estimations proved to be valuable tools for examining the models and the response variable, sorties generated. This estimation allowed the assessment of the models' performance under different starting conditions, i.e., high and low treatments, and under different trial sizes. The confidence intervals depicted in Figure 12 also help choose the correct data to use for the remainder of the statistical calculations.

Table 9

TSAR Confidence Limit Estimates

Treatment /Trials /Days	Mean Sorties Generated	Standard Deviation	Number Trials	$t_{\alpha/2}$	Lower Confidence Limit	Upper Confidence Limit	Width
H/10/30	3369.4	734.8	10	1.383	3048.040	3690.760	642.720
H/20/30	3355.2	868.2	20	1.328	3097.388	3613.012	515.624
H/30/30	3744.8	917.1	30	1.311	3525.288	3964.312	439.024
L/10/30	2920.6	254.1	10	1.383	2809.471	3031.729	222.258
L/20/30	2895.1	196.9	20	1.328	2836.631	2953.569	116.938
L/30/30	2948.7	193.0	30	1.311	2902.505	2994.895	92.390

Table 10

SORGEN Confidence Limit Estimates

Treatment /Trials /Days	Mean Sorties Generated	Standard Deviation	Number Trials	$t_{\alpha/2}$	Lower Confidence Limit	Upper Confidence Limit	Width
H/10/30	2987.2	193.4	10	1.383	2902.618	3071.782	169.164
H/20/30	2948.4	208.6	20	1.328	2886.456	3010.344	123.888
H/30/30	2974.3	177.3	30	1.311	2931.862	3016.738	84.876
L/10/30	1820.8	116.7	10	1.383	1769.762	1871.838	102.076
L/20/30	1896.3	131.1	20	1.328	1857.370	1935.230	77.860
L/30/30	1840.0	139.3	30	1.311	1806.658	1873.342	66.684

The appropriate data for the TSAR model is found by first choosing the treatment group that has the largest confidence interval; this is the TSAR high treatment group. The second decision is made by choosing the treatment within that group that has the smallest confidence interval; this is the 30-trial treatment. The treatment data to be used in the remainder of the statistical calculations for TSAR are those resulting from treatment H/30/30. The same decisions are made for SORGEN and the treatment data used for the remainder of the statistical calculation for SORGEN are also those resulting from treatment H/30/30. The assumption for the calculation of confidence intervals is that the sampled population is approximately normal; this assumption is tested next.

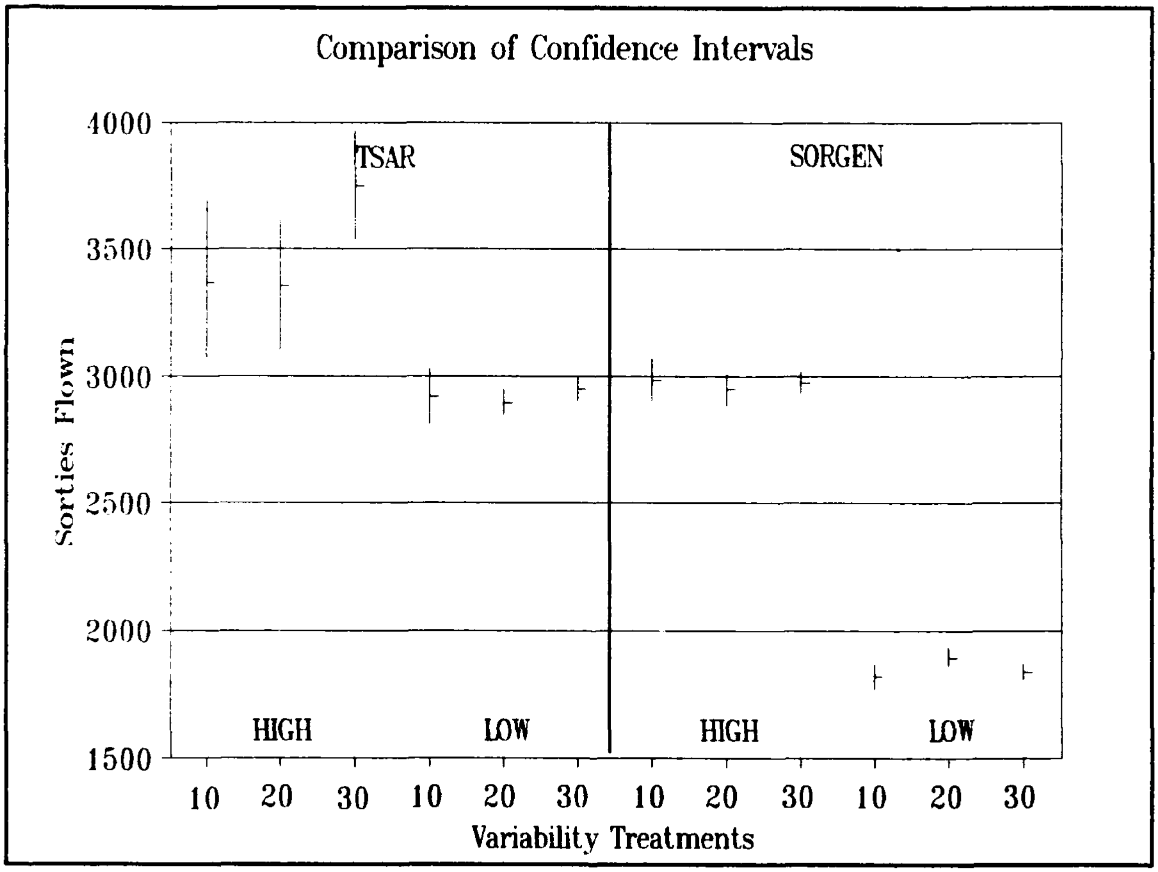

Figure 12. Comparison of Confidence Intervals

Wilk-Shapiro Test For Normality. The assumption of normality is tested using the Wilk-Shapiro Test. The end-of-trial measure of sorties flown is taken from four populations; a 30-trial high treatment for TSAR, a 30-trial low treatment for TSAR, a 30-trial high treatment for SORGEN, and a 30-trial low treatment for SORGEN. The end-of-trial statistics were submitted to SAS in a routine to measure the normality of the sampled populations. While the results vary, especially "Probability < W", the data are believed to have come from populations whose distributions are approximately normal since the values for "W:Normal", the primary measure of normality, are all close to one. The partial test results are provided in Table 11, and the complete results are provided in Appendix G.

Table 11

Test of Normality Results

Treatment	W:Normal	Probability < W	Results
TSAR H/30/30	0.953395	0.2416	Normal
TSAR L/30/30	0.918815	0.0281	Normal
SORGEN H/30/30	0.984777	0.9391	Normal
SORGEN L/30/30	0.95608	0.2824	Normal

Calculation for Number of Trials. Using the data chosen during the confidence interval estimation, the next step is to determine the number of trials necessary to achieve an 80% confidence level in the models' estimation of sorties generated. In Chapter III the confidence level was to be 80% based on the broad scope of comparison. To determine how much difference the use of 80% makes versus 95%, and to verify that 80% is reasonable, a calculation is made using 95% confidence and W = 100. The result of this calculation is that 1292.4 or 1293 trials are required to achieve a 95% confidence level with a confidence interval of 100 sorties, for TSAR; this figure would require more computer time than is economically feasible for this study and substantiates the use of a lower confidence level. A calculation was made using 80% confidence and W = 100 with a result of 552.7 or 553 trials required; this too is larger than is economically feasible. An observation made during the reasonability runs and during the variability runs showed the 30-trial runs required approximately 4.5 hours of computer time per run. This fact necessitated a smaller number of trials be conducted for each experimental run for two reasons: 1) the economics of running the models for such a long period of time per run is not feasible for the purposes of this study, and 2) the ELXSI-based system we were required to make experimental TSAR runs on was scheduled for permanent shut-down and removal. It was impossible, in terms of time available, to complete the entire series of 64 experimental TSAR runs using 30 trials per run, before the system was shut down. The test of equal variances showed there were no significant differences in variances between the 20-trial and 30-trial variability data. Based on this finding a decision is made to use 20 trials per run during the execution of the experimental runs. A final calculation is made to determine the value of W, given the use of 80% confidence and 20 trials per run. The result of this calculation is that W = 526. Based on these calculations the mean can be estimated ± 263 sorties with 80% confidence. The results of these calculations are provided in Table 12.

The trials required by SORGEN are subsequently calculated using the same formulas as those used above for TSAR. The variability is considerably less in SORGEN than in TSAR, this is evidenced in the smaller number of trials needed in

SORGEN to achieve the same confidence level as TSAR. To achieve an 80% confidence level and W = 526, SORGEN only requires 0.75 or one trial be simulated. The results of the calculations for SORGEN are provided in Table 13.

Table 12

Number of Trials Calculations for TSAR

Based on	α	$Z\alpha/2$	s	W	n
30 trials 95% confidence	.05	1.96	917.1	100	1292.4
30 trials 80% confidence	.20	1.2817	917.1	100	552.7
30 trials 80% confidence	.20	1.2817	917.1	526	19.9

A decision about the number of trials is made based on the calculations for TSAR and SORGEN, and on two outside influences: 1) the computer run-time calculations require an equal number of trials be made, and 2) the use of one trial for SORGEN may not be a sound approach to the quantitative comparison of these models. As a result, the two models are run using 20 trials per treatment during the factorial experiment. In this research two considerations drive the need to exercise the models for only 20 trials per run: 1) the economy of 20 trials, and 2) time considerations taken into account due to retirement of the host system. One final calculation is needed to verify the use of 20 trials, the test of hypothesis for equal population variances.

Table 13

Number of Trials Calculations for SORGEN

Based on	α	$Z_{\alpha/2}$	s	W	n
30 trials 95% confidence	.05	1.96	177.3	100	48.3
30 trials 80% confidence	.20	1.2817	177.3	100	20.7
30 trials 80% confidence	.20	1.2817	177.3	526	0.75

Test of Hypothesis for Equal Population Variances. The results of the confidence interval estimation in Tables 9 and 10, and Figure 12, provide a numerical and visual display of the variability in each of the treatments tested. A stronger confidence about the results of the experiment may result if the variances for the 20-trial and 30-trial data are equal. To test the variance an F-test is conducted. The test of hypothesis shows the variance of the two sampled populations are equal. The results of the tests are given in Table 14. The results of the test of variance provide further evidence that 20 trials may be used in the conduct of the experiment without significant loss in confidence about the models' estimates of sorties generated.

Conduct of the Experiment. The results of the pilot runs and the number of trials per run determination leads next to the conduct of the experimental design. The design is a 1/4 fractional factorial experiment as shown in Chapter III. Both models are run and the resulting average number of sorties flown over 20 trials are collected.

The data collected from both models are provided in the statistical analysis results obtained from SAS (see Appendix H).

Table 14

Test of Hypothesis for Equal Variance

Test Comparison	s_1 s_1^2	s_2 s_2^2	s_1 df	s_2 df	F	$F_{\alpha/2}$ $\alpha = .05$	Reject Ho
Tsar HI30 vs HI20	917.1 841,072.41	868.2 753,771.24	29	19	1.116	2.39	no
TSAR LO30 vs LO20	193 37,249.0	196.6 38,651.56	29	19	1.038	2.21	no
SORGEN HI30 vs HI20	177.3 31,435.29	208.6 43,513.96	29	19	1.384	2.21	no
SORGEN LO30 vs LO20	139.3 19,404.49	131.1 17,187.21	29	19	1.129	2.39	no

Statistical Analysis of Results. The results collected from the two models are analyzed using a SAS routine of paired differences to determine whether the models produced similar results when using equivalent databases. The results of the statistical analysis are provided in Appendix H. Also in Appendix H is the test of normality for the population of differences, required for the paired difference test.

The paired difference test calculated in SAS indicates the mean difference in sorties generated between the two models is 1074.064 and the standard deviation is 407.9631. The t-statistic calculated by SAS is 21.06198 with a P-value of 0.0001. These results indicate that the mean difference of the model treatments does not equal zero and thus Ho is rejected supporting the conclusion that the models do not produce similar quantitative results under these experimentally equivalent conditions. The results of the paired difference test are synopsized in Table 15.

Table 15

Paired Difference Test Results

Mean Difference	Standard Deviation	t-statistic	p-value	Results
1074.064	407.9631	21.06198	0.0001	Reject H_o

Of interest is the cause of the difference between the two models. To analyze the exact cause of the differences is beyond the scope of this research. However, to examine the differences between identical treatments in each model, cumulative daily sorties generated are plotted in line graphs against each other. The treatments chosen for this exercise are the 20 trial, all-high treatment from the variability tests, the all-low treatment, and three randomly selected treatments (acefg, abdefg, and bdfgh). The plots of these treatments are provided in Figures 13, 14, 15, 16 and 17 respectively. The response of each model in terms of sorties produced differs for each of the treatments. The response curves indicate the models produce different values of the measure of merit under these conditions. Of interest, however, is the extent to which the models differ when the trial length is extended to some value longer than 30 days. Some of the response curves appear to begin converging toward the end of the 30-day period for which these models were tested see Figures 13, 15, and 16. The possibility exists that both models are valid tools for their intended purposes. The

TSAR model is a high fidelity model and provides a high level of detail to the user; its use may be more appropriate in simulations of relatively short duration such as this 30-day period. The SORGEN model, however, may be equivalent in its production of sorties past the 30-day period and may prove to be the most appropriate choice of models for simulations of greater than 30 days duration. The impression left by these graphs is that the overall response of both models is similar. The cause of the differences and the sensitivity to factors and treatments is left to future research, because this analysis will require an in-depth exploration of the models' code and algorithms.

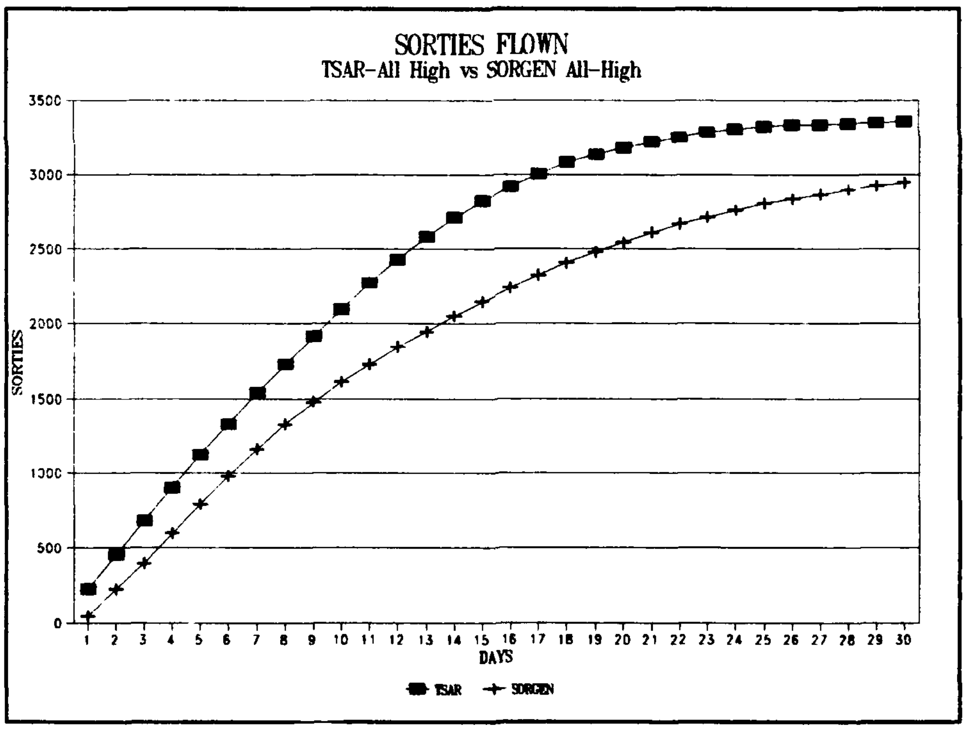

Figure 13. Sorties Flown Comparison All High Factors

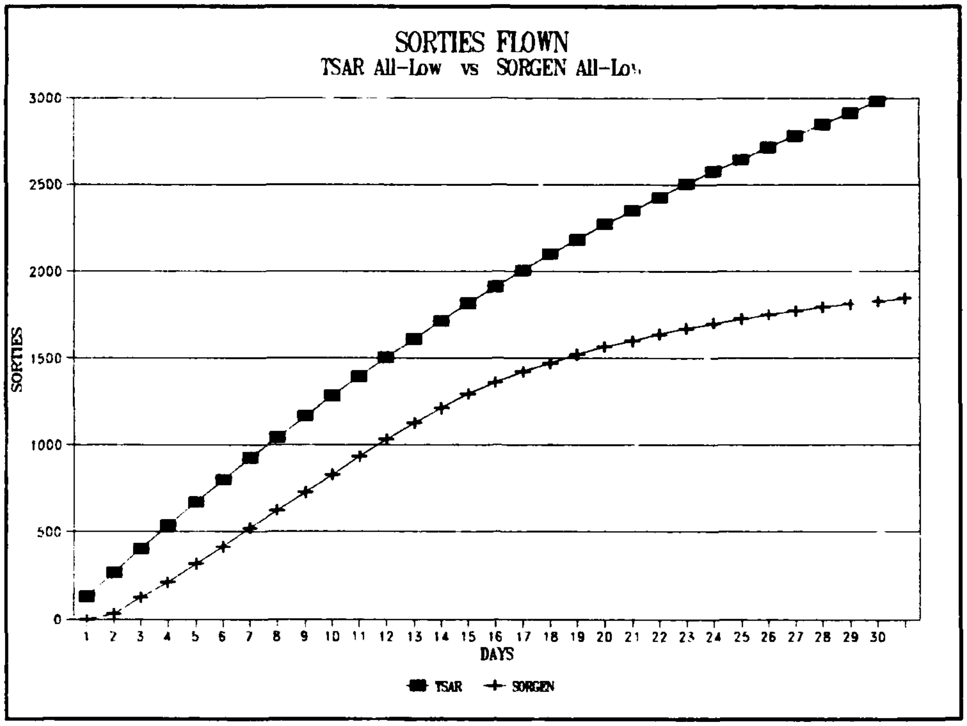

Figure 14. Sorties Flown Comparison, All Factors Low (Treatment 1)

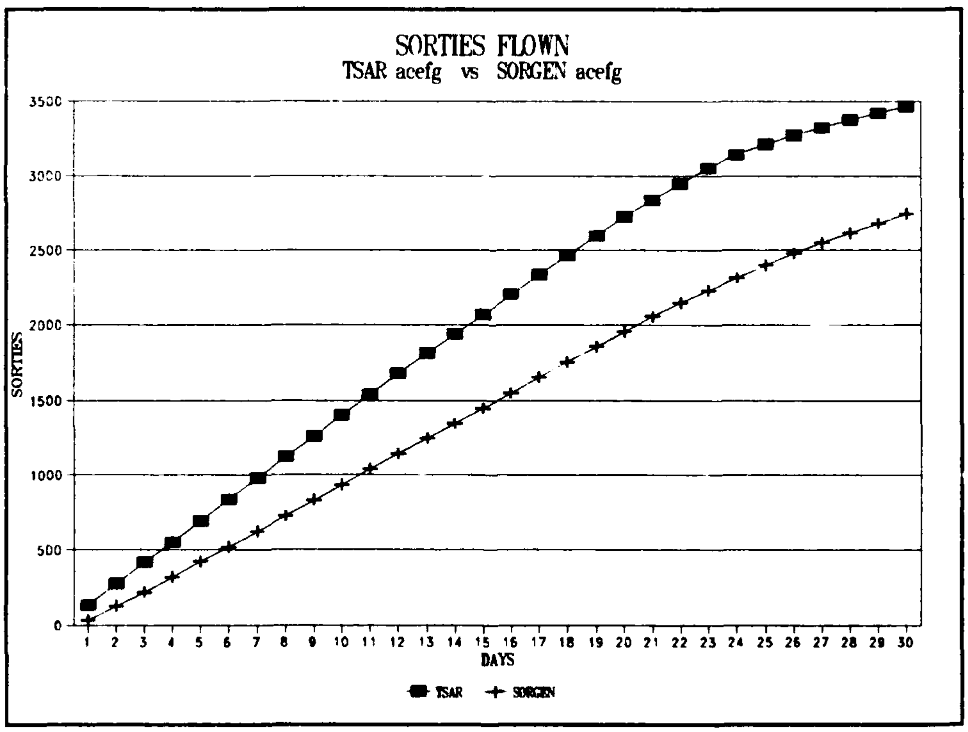

Figure 15. Sorties Flown Comparison Treatment 12

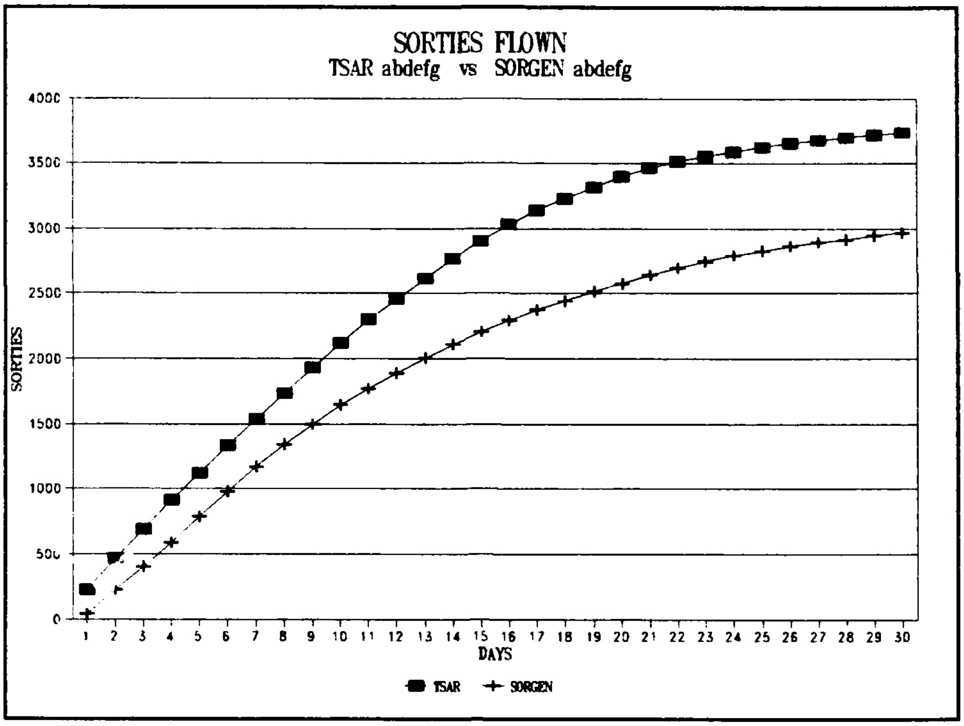

Figure 16. Sorties Flown Comparison Treatment 29

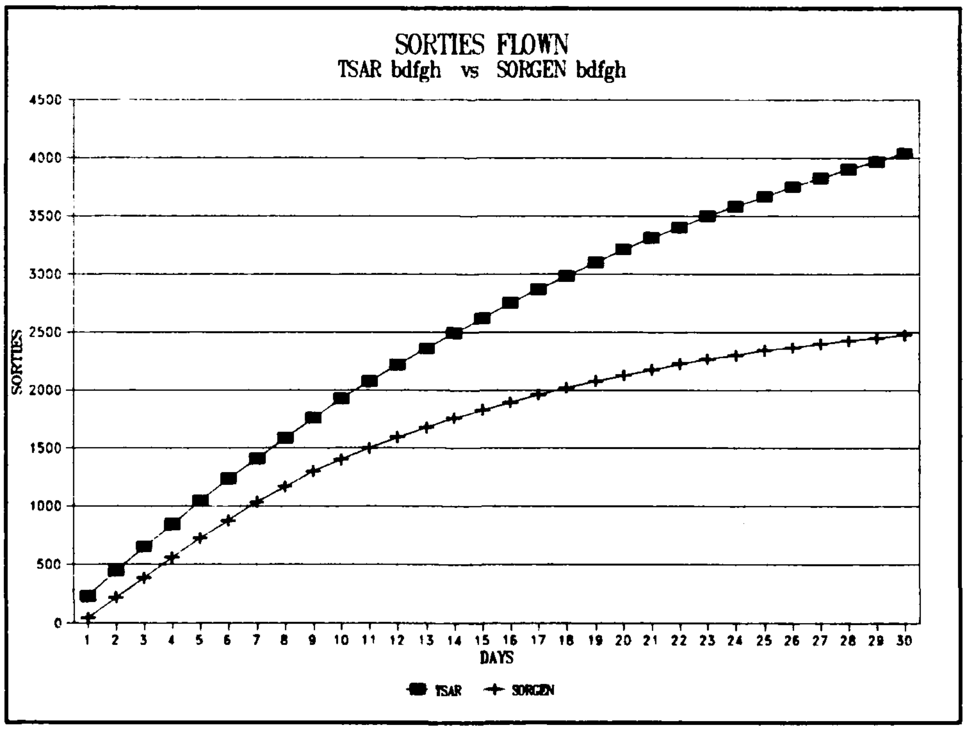

Figure 17. Sorties Flown Comparison Treatment 54

Test of Normality for Differences. To assure the paired difference test is appropriate it should meet the assumptions of the paired difference test. The Wilk-Shapiro test for normality is used and the results, W:Normal = 0.98262, Prob<W = 0.7669, and p-value = 0.0001, indicate the differences come from a normally distributed population.

Synthesis and Summary of the Qualitative and Quantitative Comparisons

The synthesis of the three qualitative components and the single quantitative component is one based on individual or team impressions; evaluated in terms of the models under study and the questions to be addressed by the models. The TSAR and AMTAF models demonstrate a reasonable amount of overlap in the circles found in Figure 5.

Figure 5 illustrates that documentation is the foundation of the model comparison methodology and is likely the most important single aspect of a simulation model. Documentation is a notable weakness for the SORGEN model. The effects of documentation weaknesses ripple throughout the use of the model.

The databases and features of each model have their own strengths and weaknesses. The SORGEN database is generally easier to learn and manipulate due to the VICS database management interface. The TSAR database simulates more features of the airbase environment, the majority of which are depicted in Figure 10, and this makes the model more complex and difficult to use.

There are useability aspects scattered throughout the models that impact other components of the model comparison construct. Both models' useability suffer because of their lack of documented debugging tools. The impact of adequate documentation on the efficient and effective use of a model is again seen.

There may be some impact from each of the three qualitative components of the model comparison construct that affect the quantitative nature of the models. A statistically significant difference was observed in the response variable used for quantitative comparison, although there is no evidence to indicate why this occurred. An assumption must be made at this point that differences exist in the model code, algorithms, and possibly model logic that drive the differences in the quantitative results of these models. The answer to this topic of interest will only come with further research.

The methodology developed in this research is that of a model comparison construct. The illustration of Figure 5 urges the researcher to investigate not only the individual components found within each circle of the figure but also the interaction among the figures. This research does not attempt to capture all the possible interactions and leaves this to the individual who chooses to use the methodology. The purpose of the comparison of models is to provide a structured process for identifying the similarities and differences between two or more models, making decisions about the utility a model may hold, and establishing the confidence necessary to accredit a model for specific purposes within the decision-maker's organization.

The framework established in Chapter III is refined in Chapter V, using knowledge gained through the course of this study. In addition, recommendations for future research in the area of airbase operability and model comparison are provided.

V. *Conclusion and Recommendations*

Overview

This research represents the development of an alternative methodology for the comparison of models, with the objective of establishing credibility in computer simulation models. The traditional, and still optimum approach, is to verify and validate a model using methods developed by authorities like Balci, Banks and Carson, Law and Kelton, Shannon, VanHorn, and others. The methods sometimes require knowledge and resources that may not be available in some organizations, but the absence of such knowledge should not be an absolute deterrent to using new, untested models. These reasons led to the belief that a comprehensive alternative could be developed that would allow decision-makers to develop confidence in new models by comparing them against models already accredited by decision-makers within the organization.

The Sortie Generation (SORGEN) module of the All Mobile Tactical Air Force (AMTAF) model presents an opportunity to test the new methodology developed in this research. The SORGEN model is a relatively new model, developed in the late 1980s, to simulate the operability functions of our airbases. The model was procured by the Mission Area Planning section of Aeronautical Systems Center (ASC/XRS) to improve their mission area planning capability. Unfortunately, the model does not have an extended user group purportedly due to lack of model credibility. The most closely related model in wide use is the Theater Simulation of Airbase Resources (TSAR) model which is used by various organizations for planning and analysis, including the Air Force Center for Studies and Analysis (AFCSA) and the Munitions Development Branch of Aeronautical Systems Center (ASC/YQ). The credibility of SORGEN may be established if it compares favorably with a widely accepted model like TSAR.

This research in no way attempts to verify or validate either model. In fact, these researchers are not aware of any formal attempt to validate the TSAR model, and must assume the developers did conduct appropriate verification. The TSAR model has simply been used over a broad range of variables and by a wide user group and has, by default, established its credibility with analysts and decision-makers within the Air Force. This de facto credibility, however, should not be slighted because it is based on the model users' and decision-makers' expert opinion the model's output provides a reasonable representation of the real systems it simulates. As a baseline for comparison the TSAR model provides an excellent point of departure.

The remainder of this chapter summarizes the research accomplished in this study. First, the questions that prompted the study and our hypotheses for the study are revisited and answered. Second, a review of the stated methodology is provided along with the researchers' view of its strengths, weaknesses, and needed improvements. Third, recommendations for future research are made, and finally, a summary of the research is presented.

Summary of Research Questions and Hypotheses

This research is based on a series of questions and hypotheses that include management, research and investigative questions, and hypotheses about the equivalence of the models under study. These questions are addressed in the research, in the order they are given above. To provide a summation of the research, however, they are presented in reverse order, since to answer the higher order questions, one must first start with the more basic ones. The discussion first covers the hypotheses, second the investigative questions, third the research questions, and finally the management questions.

The Research Hypotheses. This research is intended to demonstrate whether the SORGEN and TSAR models are sufficiently equivalent, both qualitatively and quantitatively, so that SORGEN's credibility may be established among its potential users and the users of TSAR. The hypotheses proposed were:

1. H_O: $SORGEN_{QUALITATIVELY} = TSAR_{QUALITATIVELY}$

   H_A: $SORGEN_{QUALITATIVELY} \neq TSAR_{QUALITATIVELY}$

2. H_O: $SORGEN_{QUANTITATIVELY} = TSAR_{QUANTITATIVELY}$

   H_A: $SORGEN_{QUANTITATIVELY} \neq TSAR_{QUANTITATIVELY}$

The models were shown in this research to contain a wide variety of differences. For the conditions under which the two models were run, those differences are probably large enough to call the models non-equivalent. The conditions under which these models were run, however, may be the key to why the models do not produce similar results. The graphs presented in Chapter IV, in Figures 13, 14, 15, 16, and 17, lead to other questions about the models' response under the given conditions of this study. In several cases, Figures 13, 15, and 16, the response curves appear to be converging at the end of 30 days. Under the conditions of 60 or 90 days the difference seen in these graphs might be insignificant at the end of those time periods. This research cannot say conclusively that the models are non-equivalent, only that under the conditions of this research the models do not compare favorably (see Table 16). Each model may be suitable for research and applications analysis under different conditions. This research began by using a TSAR-type scenario and fitting an AMTAF database to it. If the scenario is reversed and TSAR is fitted to an AMTAF-type scenario the reverse observation may be observed.

The methodology, however, achieved its purpose, to systematically explore the models and their documentation and to document the differences found. This forms the basis on which the model user or decision-maker can begin to build some confidence that a model provides useful and useable information.

Table 16

Experimental Conditions

Condition	Scenario/Setting
Number of Bases	One
Initial Aircraft	72
Period of Simulation	30 Days
Resupply	Fuel and Munitions Only
Attack	None
Aircraft Attrition Rate	Day 1: 1.2%, Day 2: 1.1%, Day 3 - 30: 1.0%
Aircraft Battle Damage	None
Aircraft Ground Damage	None
No./Type Missions	5 Air-to-Air
Munition Use Rate/Mission	100%
Baseline Database	Modified TSAR F-15C Database

The Investigative Questions. Five investigative questions were proposed in order to focus the investigation of whether SORGEN and TSAR are qualitatively and quantitatively equivalent. To what extent are the models equivalent with respect to:

1. The general classification and level of performance?

2. The input requirements and characteristics?

3. The output data format and characteristics?

4. The man-machine interface (ease of use)?

5. Output data given equivalent input?

Each of these questions forms the basis of the methodology developed in this research and is addressed in detail in the findings of Chapter IV. These questions guide the researcher through important facets of the models and in doing so focus the study in a way that allows the researcher to find and examine crucial differences between two models (see Table 17). The findings in this study reasonably support the use of model comparison, both qualitative and quantitative, as a method for establishing model accreditation and acceptance. There was no expectation that the models would be identical; similarities and differences do exist. The differences found in this study only allow an opinion to be drawn about the models under one set of experimental conditions. Experimentation under other conditions is necessary to draw further conclusions about the depth and breadth of similarities and differences between TSAR and SORGEN. The repeated use of the methodology with strong scientific experimentation will allow the analyst and decision-maker to form stronger opinions about the models under conditions formed to meet their needs.

Table 17

Investigative Summary

Comparison Category	TSAR	SORGEN
* Classification	Airbase Operability Model	Sortie Generation Model
* Level of Performance	High fidelity, theater of airbases	Medium fidelity, theater of airbases
* Input Data Requirements	Varies from small to large (scenario dependent)	Varies from small to large (scenario dependent)
* Input Data Format	80 Column IBM card, 117 card types 15 categories	Relational Database, 10 types w/ multiple relations
* Output Data	User controlled w/event and cumulative statistics	User controlled w/event and cumulative statistics
* Output Data Format	Preformatted, data-only, & post processing capability	Preformatted only
* Man-machine Interface	Cumbersome, poorly documented debug capability, long run times, parallel run capability	Semi-friendly, poorly documented debug capability, shorter run times, batched (single run) capability
** Output Given Equivalent Input (Sorties Estimated)	Higher for 30 day scenario	Lower for 30 day scenario
** Output Variability (Sorties Estimated)	High, requires large number of runs for high confidence intervals	Low, requires smaller number of runs for high confidence intervals
* Qualitative Measure	** Quantitative Measure	

The Research Question. The model comparison methodology and the five investigative questions provide the answer to the research question, "What constitutes a sound model assessment methodology?" Using the research methodology applied in this work the researchers revealed numerous qualitative differences and a quantitative difference based on the statistical test of one response variable, sorties generated. This study indicates the methodology is sufficiently detailed to allow a research team to systematically identify important differences between models. The methodology, however, must necessarily be an iterative process. This study is in no way comprehensive enough to identify all, or even the most important differences of the two models. Further research is needed under different conditions to learn more about the models.

Model comparison appears to be a sound model assessment methodology and a means for establishing the credibility of models, when applied rigorously and objectively. It allows the analyst or decision-maker to simultaneously evaluate the similarities and differences, and assess the strengths and weaknesses of both models. The iterative and objective application of model comparison under various circumstances provides important information about the response of the multiple models under identical operating conditions.

The Management Questions. This research began recognizing that TSAR enjoys a higher level of use within the Air Force analysis community than does SORGEN, even though SORGEN is described as having similar capabilities. The need to answer the management questions, "To what extent are the two models similar?" and "How can AMTAF and TSAR be compared to determine the extent to which they are equivalent?" provided the impetus to derive a technique that compared the simulation models more dynamically than previously attempted. The question of how to compare the models was answered in the development of a methodology that compares simulation models from both a qualitative and quantitative perspective. Conducting the proposed qualitative comparison produced findings that indicate there is a high degree of similarity in the two models. However, the findings of the quantitative comparison indicate the models are not equivalent under the conditions used in this research. The extent to which the models are quantitatively equivalent remains to be adequately addressed for several reasons. First, the translation of a TSAR database for use in SORGEN may have weighted this comparison in TSAR's favor. Translation of a SORGEN database for use in TSAR is suggested for a subsequent comparison. The results of additional comparisons are needed to support or disprove the significance of these findings. Second, limiting the period of the simulation to 30 days does not fully evaluate the limits of either model. The similar shape of the sampled response curves for cumulative sorties flown suggests the need to compare the performance for a longer period to see if there is a point of convergence. Finally, constraining the environment to a no attack case with no aircraft battle or ground damage and no outside resupply of parts, support equipment and personnel removes a significant amount of potential for variability that may have a normalizing affect on the two models predictions. Should future studies determine and bound regions of quantitative equivalence, a basis for the acceptance of SORGEN will be realized solely from comparability with TSAR. Assuming that subsequent studies will use the methodology developed here makes necessary a summary analysis of its strengths, weaknesses, and needed improvements, the topic of the next section.

Analysis of the Methodology

The original intent of this research was to develop and test a methodology for comparing two simulation models, both qualitatively and quantitatively. The construct first proposed in Chapter III is used throughout the study to frame our efforts. Figure 18 presents again the basic construct for the purpose of discussing the strengths and weaknesses discovered as the research progressed. The discussion begins with a brief overview, followed by the strengths of the methodology and its weaknesses. The analysis is concluded with a proposed change to the methodology to overcome a portion of the weaknesses discovered.

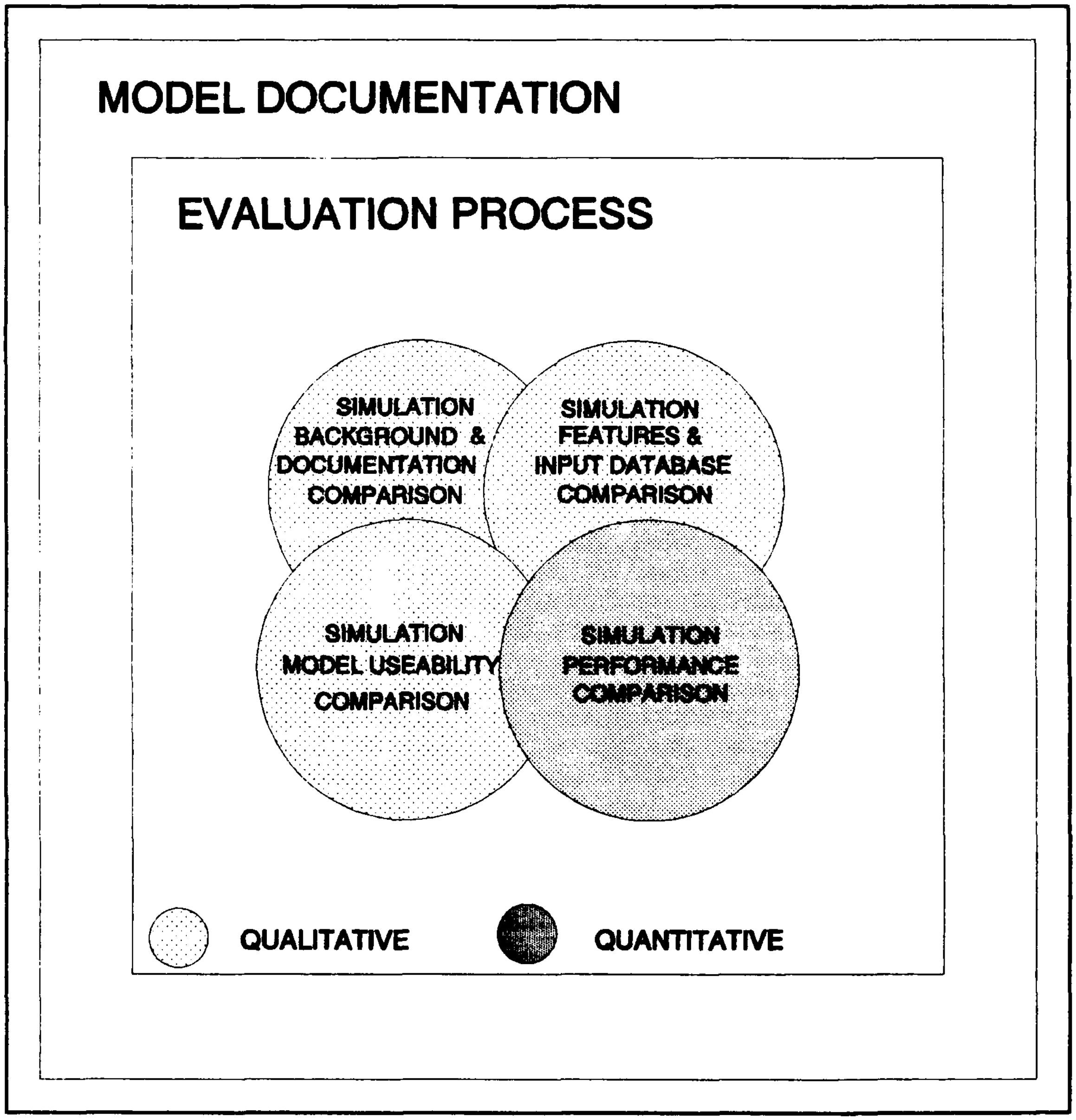

Figure 18. Construct for Model Comparison

Overview. The qualitative and quantitative comparison methodology designed for this research is intended to provide a more complete comparison of two models than attempted in previous efforts. The qualitative comparison is believed necessary to provide potential users a comparison of the traits and characteristics of two simulation models with respect to the environment they are intended to simulate.

The qualitative comparison of input and output characteristics enables potential users to assess the suitability of the models to his/her needs. Knowledge of what is required to run the model and what outputs are produced, provide useful information that can be used to determine the suitability of the models for a particular purpose. In the conduct of the research, strengths and weaknesses in the proposed methodology were observed. The discussion is first turned to the apparent strengths.

Strengths. The investigation of the models' background, documentation, and features provides a sound basis from which to base a one-to-one model comparison. In the case of TSAR and SORGEN, the two models proved dissimilar in the qualitative aspect. The quantitative comparison of the models, exercised the two simulations more thoroughly than were earlier model comparisons. The selection of multiple experimental factors provides a sound basis for a broad comparison of the models' predictions given equivalent inputs. Here too, TSAR and SORGEN proved to be dissimilar. While the qualitative and quantitative comparisons produce common outcomes, taken together they provide a more encompassing framework on which to determine how and when to use a particular simulation. It also provides a means of estimating how the outcomes differ. Additional analysis could provide a comparison of model sensitivities to the input factors, an aspect of comparison left to future research. The comparison methodology was not without its weaknesses, which are discussed in the next section.

Weaknesses. Validation is an important aspect of modeling, but full validity of a stochastic Monte Carlo model is not likely since a model is never a perfect representation of the environment it represents. The methodology used in this research does not validate a model. Even if the comparison of two models includes a model that is fully validated, the methodology would be insufficient to fully validate the second. The methodology, in this case, would only suggest the credibility of the second model based on the fully validated model. To achieve validation of the second model, a comprehensive validation effort would have to take place to compare it against the environment it is intended to represent.

Another weakness, one which directly influenced the progress of this research, is that the proposed methodology does not assess the models' database capacity. At a critical time in the course of this investigation changes were necessary in order to use the SORGEN model. The database translated for use in SORGEN was too large for the initial configuration of SORGEN's memory arrays. The arrays were resized four times during the conduct of the research, with the final limits roughly eight times the original configuration. This particular problem is believed common, especially in the early stages of the model's life cycle. This aspect of the comparison should be included in future studies to preclude expending time, effort, and money on models, if the model support environment is unable to reconcile the problem. The model architecture comparison is added to the features and input databases portion of the model comparison construct as illustrated in Figure 19. The model architecture comparison should include a review of programmer/analyst manuals, and include questioning the model support staff about the ability of the model in question to handle databases in the size range anticipated by the users.

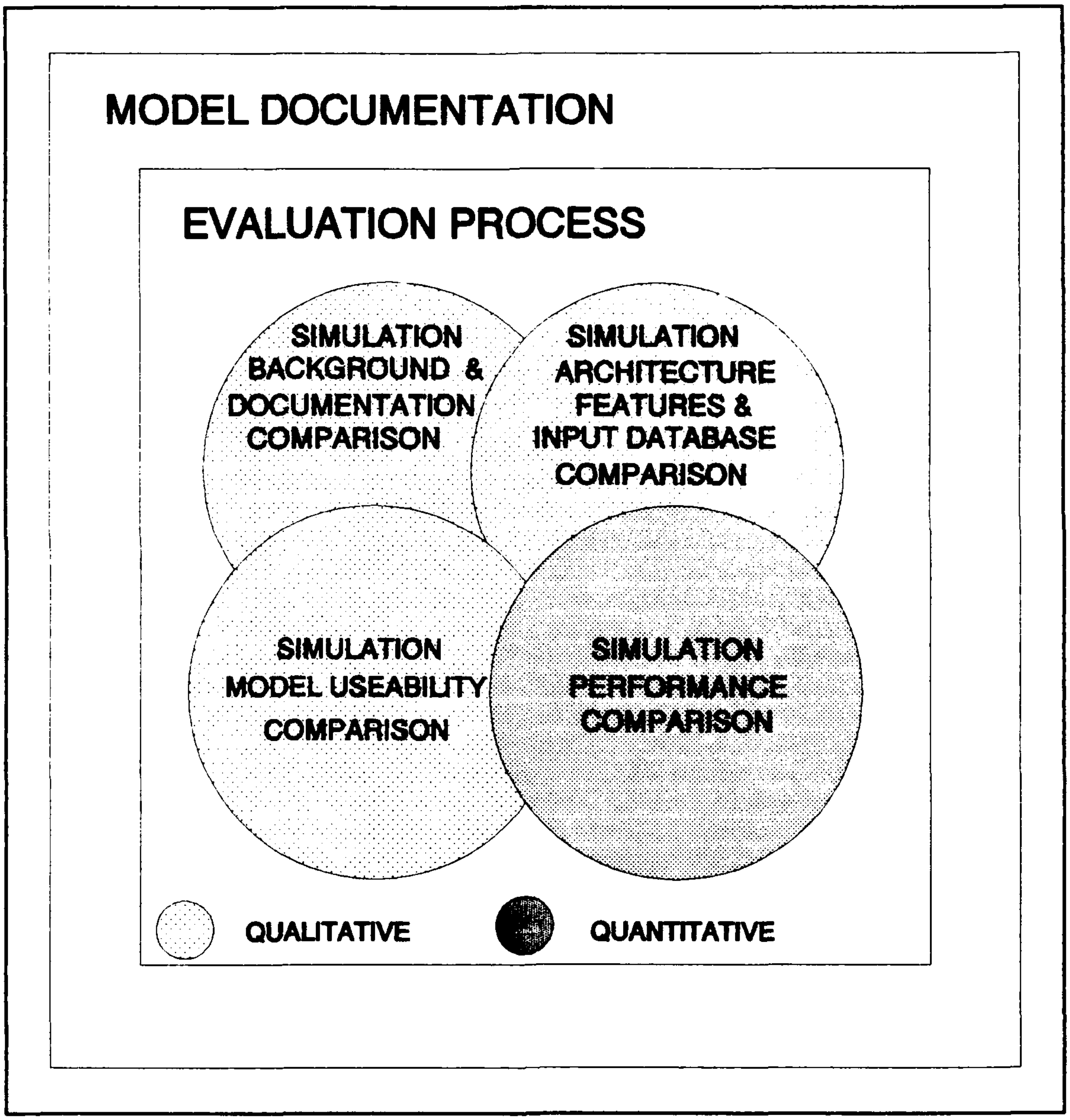

Figure 19. Improved Model Comparison Construct

Through the course of this research, areas of potential future studies presented themselves. The methodology proposed in this study has now been tested and modified to strengthen it. It is now ready to be subjected to further rigorous testing and modification as necessary. Recommendations for future research are presented in the final sections of this chapter.

Recommendations for Future Research

Numerous areas of possible research were noted during the accomplishment of this effort. The most notable areas are enumerated in the following paragraphs.

Including New Features. The comparison of TSAR and SORGEN in this research is done without exercising the attack features offered by TSARINA, the companion threat model common to both TSAR and SORGEN. The inclusion of an attack feature would introduce more variability to the models and further stress the airbase environment in its effort to generate sorties. Future research should qualitatively and quantitatively compare the two models in an attack scenario.

This research, while more encompassing than earlier studies, is limited to a few model features. Future research should expand the logistics analysis to include outside resupply of parts and part level computations. This would require the comparison of AMTAF's LOGSIM module which simulates parts resupply and resource level computations for use by SORGEN.

Sensitivity Analysis. Since this research finds SORGEN and TSAR similar in purpose but different qualitatively and quantitatively, future research might focus on determining how SORGEN and TSAR differ quantitatively. This could possibly be achieved by testing the sensitivity of each model to the input factor levels. This research used a 1/4 fractional factorial experimental design that allows all the two-way interactions to be measured. Using the data from this study, a statistical analysis of the data could determine which factors and their levels provide the greatest impact to the production of sorties in these models.

This study uses a trial length of 30 days to assess the airbase environment and the sortie generation capability of both models. One of the findings of Chapter IV suggests that with a longer trial length, the quantitative differences between TSAR and SORGEN might be insignificant. Future research might reproduce this study, changing the trial length to 60, 90, or 120 days, to examine the response of both models under these new conditions.

Future research could focus on the analysis of the data produced in the TSAR/SORGEN comparison. While we used cumulative sorties over 30 days as a single measure of model performance, analysis of the data using daily sorties flown would facilitate the analysis of differences in model performance.

Reverse Comparison. This study compares SORGEN to TSAR by translating a TSAR database for use by SORGEN. A second comparison, where TSAR is compared to SORGEN, should be accomplished; translating a SORGEN database for use by TSAR.

Methodology Testing. The comparison of a third model to SORGEN and TSAR is yet another area where future research could focus. Translation of the databases developed for this research, for use in the Combat Base Assessment Model (CBAM), produced by GARJAK Research Inc., is one option.

Summary

This research illustrates a methodology for establishing de facto accreditation by comparing simulation models. The purpose of the model comparison methodology is to provide another alternative for establishing a model's credibility for specific decision-making purposes. It represents a reasonable method for users and decision-makers to establish the confidence necessary to use models for decision-making. When its application is both rigorous and iterative, it yields a relatively comprehensive review of the models' attributes. It does not yield any measure of validity; only a stronger measure of credibility and acceptance are obtained through the use of model comparison as presented in this study. The strongest measure of a model's usefulness, however, may be the decision-makers' ultimate acceptance or accreditation of the model and its eventual use for decision-making. Finally, the ultimate utility of the methodology may only be realized through its continued use and modification to fit the needs of individual studies.

Appendix A: Factors, Treatments, and Random Numbers

Factor Assignments

The following factors are those used in both TSAR and AMTAF.

a - Aircraft

b - POL (fuel)

c - Munitions

d - Mission

e - Personnel

f - Spares (parts)

g - AIS (avionics intermediate shops)

h - Support Equipment

Reasonableness Array Design

Treatments	Random Number
(1)	797867
abcdefgh	90195
abcdefg	379461
abcdefh	49631
abcdegh	300067
abcdfgh	76685
abcefgh	212505
abdefgh	569349
acdefgh	681941
bcdefgh	675455
h	73185
g	872111
f	147789
e	82625
d	505187
c	863507
b	326649
a	632517

Variability Array Design

Treatments	Random Number
H/10/30	700021
H/20/30	444839
H/30/30	840825
L/10/30	119961
L/20/30	213521
L/30/30	678669

Experimental Array Design

Treatments	Random Number	Treatments	Random Number
(1)	703721	eg	259123
abcfgh	764291	adfgh	388922
bcdeg	903529	bcd	59127
adefh	532201	abdh	652659
efgh	284605	cdfg	663097
abce	792657	acegh	62139
bcdfh	85885	bef	779801
adg	730418	abdefg	403583
cdgh	245865	cdeh	95989
abdf	453685	acf	251313
beh	99015	bgh	538141
acefg	488809	bdefh	872029
cdef	80155	acdeg	267249
abdegh	865427	cfgh	324825
bfg	17251	ab	139195
ach	133351	bdg	75753
abcg	844985	acdfh	91771
fh	414843	ce	286927
ade	276613	abefgh	155075
bcdefgh	585995	bcefg	216807
abcefh	368319	aeh	256997

Experimental Array Design (continued)

Treatments	Random Number	Treatments	Random Number
df	142061	defg	58015
abcdgh	610343		
bch	662789		
afg	86165		
degh	70653		
abcdef	582475		
acdefgh	173617		
bde	50425		
abfh	341555		
cg	882867		
acd	906321		
bdfgh	85705		
abeg	863203		
cefh	208757		
aef	631937		
bcegh	20045		
abcdfg	65213		
dh	132987		
agh	792763		
bcf	97063		
abcdeh	190503		

Appendix B: TSAR and SORGEN Feature Comparison

The comparison of simulation features is presented in this appendix. Explanatory notes are presented at the end of the appendix to clarify differences in the implementation of features. The purpose of this appendix is to provide a comprehensive listing of model features, however, the complexity of TSAR and the documentation weaknesses of the AMTAF modules may preclude completely achieving this goal. The reader is advised to consult the model documentation and/or experienced users if a desired feature is not found. The absence of some features in SORGEN may be overcome by experienced modelers using alternative implementations. The investigators are constrained to documented features and implementations for the purpose of this research. In the following table, the column headed "Simulation Feature" lists a brief title of the capability, with explanation where needed. The other two columns headed "TSAR" and "SORGEN" indicate the presence or absence of the feature in the respective model. There are nine features where a broader evaluation is needed than provided by yes or no. In these instances the notation, #n, refers to the footnoted comments that provide the needed explanations.

TSAR and SORGEN Simulation Features Comparison

Simulation Feature	TSAR	SOR GEN
AGE (aerospace ground equipment) for off equipment maintenance partially capable	yes	no
AGE for off equipment maintenance fails and requires repairs	yes	yes
AGE for on equipment maintenance fails and requires repairs	yes	yes
Air abort	yes	no
Air traffic control	yes	no
Aircraft attrition	yes	yes
Aircraft battle damage	yes	yes
Aircraft battle damage repair	yes	yes
Aircraft decontamination (postflight)	yes	yes
Aircraft ground damage repair	yes	yes
Aircraft ground damage	yes	yes
Aircraft mission prep	yes	yes
Aircraft off equipment maintenance	yes	yes
Aircraft on equipment maintenance	yes	yes
Aircraft rescheduled and reconfigured	yes	#1
Aircrews	yes	#2
Alert aircraft	yes	no
Alternative off equipment repair procedures	yes	no
Alternative on equipment repair procedures	yes	no
Alternative weapon loads when primary munitions not available	yes	no
Alternative weapons loads per mission	9	0
ATC cancellation	yes	no
Attack damage to resources	yes	yes
Attrition modifiers	yes	yes
Base level self sufficiency (BLSS)	yes	no
Base repair from unique procedures	yes	yes
Base repair from randomly selected alternative procedures when short of primary people/materials	yes	no
Basic munitions loads	yes	#3

TSAR and SORGEN Simulation Features Comparison (continued)

Simulation Feature	TSAR	SOR GEN
Break rate modifiers	yes	#4
Building materials for runway and shelter repair	yes	no
Cannibalization of parts from aircraft to fix other aircraft	yes	no
Cannibalization of SRUs from LRUs awaiting repair	yes	yes
Cannibalized parts possibly broken	yes	no
CE (civil engineering) personnel	yes	no
Central supply reorder	yes	yes
Centralized theater repair facility	yes	yes
Check flight following specific maintenance tasks	yes	no
Chemical attack	yes	no
COMO (combat oriented maintenance organization) maintenance structure	yes	no
Condemnation of parts	yes	yes
Conventional attack	yes	yes
Cross trained personnel	yes	yes
Cross training level of proficiency (fully qualified versus task assist)	yes	no
Cross training specified by individual task	yes	no
Deferred maintenance tasks	yes	yes
Depot maintenance	yes	yes
Diversion due to runway closure	yes	yes
DOB (dispersed operating base)	yes	yes
Early morning inspection	yes	no
Facilities	yes	yes
Facilities (chemical features)	yes	no
Facility repairs (shelters)	yes	no
Filler aircraft	yes	yes
Flight scheduling	yes	yes
Flight scrambling	yes	no
Ground abort	yes	yes

TSAR and SORGEN Simulation Features Comparison (continued)

Simulation Feature	TSAR	SOR GEN
Ground personnel for aircraft maintenance	yes	yes
Hospitalization time	yes	no
Hot pit refueling after landing	yes	yes
Independent bases	yes	yes
Initial supply stocks	yes	yes
IPE (individual protective equipment)	yes	no
Late launch	yes	yes
Lateral supply MOB <=> DOB	yes	/5
Management policy simulations	yes	yes
Manning levels	yes	yes
Mission assignment by aircraft condition	yes	no
Mission dependent munitions loads	yes	yes
MOB (main operating base)	yes	yes
MOS (minimum operating surface) repair selection algorithm	yes	yes
Multiple aircraft types assigned to a base	yes	yes
Munition use rate controllable	yes	no
Munitions	yes	yes
Munitions assembly	yes	no
Munitions load effectivity rating	yes	no
NRTS (not repairable this station) of parts	yes	yes
Number of shops for ground personnel and support tasks	25	unl
Part cost accounting	yes	yes
Part initialization per AFM 67-1 policies (LOGSIM calculations based on level and demand)	yes	no
Parts (LRU/SRU/Bit and Piece)	yes	yes
Parts repaired at base (SRU)	yes	no
Parts repaired at base (LRUs with SRU consumption)	yes	yes
Parts repaired without SRU consumption (bench stocked parts or adjustments etc)	yes	yes
Passive defenses (use of shelters and tasks with shelter doors closed)	yes	no

TSAR and SORGEN Simulation Features Comparison (continued)

Simulation Feature	TSAR	SOR GEN
Peacetime operating stocks	yes	no
Phase deferred	yes	no
Phase inspection done at night	yes	no
Phase inspection (scheduled maintenance)	yes	yes
POL (petroleum, oils, and lubricants)	yes	yes
Pooled resources (personnel)	yes	yes
Postflight inspection	yes	yes
Repair, multiple step procedures	yes	yes
Repair, priorities	yes	yes
Repair, single step procedures	yes	yes
Replacement of filler aircraft from CONUS (continental United States)	yes	#6
Resource replacements, ordered from CONUS	yes	yes
Resources resupplied (POL, munitions, personnel, parts, etc)	yes	yes
Resupply, theater	yes	yes
Runway crater repair specific to crater size	yes	no
Runway repair	yes	#7
Runways	yes	yes
Salvage parts from aircraft with non-repairable damage	yes	no
Scheduled maintenance	yes	yes
Shelter repair after attack	yes	no
Shelters	yes	yes
Shipment priorities	yes	yes
Single base	yes	yes
Sortie allocation to DOB when MOB runway closed	yes	yes
Sortie demand	yes	yes
Speed up procedures (on/off equipment tasks)	yes	no
SRU (shop replaceable unit) repair	yes	yes
Support equipment resources	yes	yes

TSAR and SORGEN Simulation Features Comparison (continued)

Simulation Feature	TSAR	SOR GEN
Task alternatives (less people more time etc)	yes	no
Task expediting (speed up procedures for on/off equipment tasks and preflights)	yes	no
Taxiway repair algorithm	yes	no
Taxiways	yes	yes
Temperature considerations for personnel using chemical protective equipment	yes	no
Theater reporting system	yes	yes
Theater wide resource management	yes	#8
Through flight inspection	yes	no
Time delay following attack	yes	yes
Transportation of theater resources	yes	yes
TRAP	yes	yes
TRAP (tanks, racks, adapters and pylons) tracked by aircraft	yes	no
Unexploded ordinance removal following attack	yes	#9
Unscheduled maintenance	yes	yes
User defined maintenance procedures	yes	yes
War readiness spares kit (WRSK)	yes	no

Notes

1. SORGEN does scrub missions but the aircraft is returned to the ready pool minus trap and munitions to be re-prepped for another mission.

2. SORGEN does not model aircrews directly, however, it would be possible to define aircrews and assign them to a task associated with the flight time. There is no known mechanism to eliminate an aircrew under this method if the aircraft were to be lost to attrition.

3. Basic munitions loads are not modeled directly. It could be made a scheduled maintenance task and munitions consumed using the LRU/Consumables feature in the task definition.

4. TSAR allows the break rates to be modified by subsystem by varying the rate by shop type. SORGEN permits only an overall change in break rate.

5. SORGEN permits lateral supply only from MOB to DOB while TSAR permits two-way lateral supply among a specific subset of bases.

6. CONUS resupply of assets for SORGEN is handled via LOGSIM.

7. SORGEN models runway repair as a time delay equal to the number of craters to be repaired times a crater repair time. It does model parallel runway repairs based on the user specified number of parallel repairs. No personnel, equipment, or materials are simulated.

8. LOGSIM models theater wide resource management.

9. SORGEN models unexploded ordinance removal as a time delay in conjunction with post-attack runway survey. In neither case are personnel or procedures modeled.

Appendix C: Significant Differences Between TSAR and SORGEN

Air abort - In the operation of an airbase, the probability of aircraft returning from a mission early, without conducting the required operations, is distinct. The ability of a simulation model to consider this aspect of reality is believed important in estimating the ability to generate and sustain sortie rates. This feature is present in TSAR but not in SORGEN. SORGEN does have a sortie effectiveness estimator, the function of which is not clearly defined in the documentation, which may provide some capability in this regard.

Alert aircraft - During periods of increased readiness, Air Force operations frequently use alert or ready aircraft to meet short notice mission requirements. This taxes the airbase logistics infrastructure since aircraft are taken from the resource pool, prepared for a specific mission, weapons loaded, and placed in a ready-to-launch status. However, these aircraft are no longer available to meet routine mission requirements. TSAR has the ability to model this feature, SORGEN does not. Available aircraft could be reduced in SORGEN through a modification to base resources using the BASE_MODS database simulating the reduction of available aircraft due to alert status. Aircraft could be restored in a similar manner via the filler aircraft feature. There is still no means to simulate short notice launch and recovery brought on by an alert posture.

Alternative procedures - TSAR permits the identification of alternative maintenance procedures in situations where parts, personnel, or support equipment are unavailable. This provides a degree of realism representative of how tasks would be achieved in a wartime scenario. SORGEN has no similar capability.

Alternative weapons loads - When primary munitions are not available to load an aircraft tasked with a specific sortie, TSAR permits the identification of nine alternative weapons loads which can be used with a lesser degree of effectiveness. If the primary munitions are not available in SORGEN, the aircraft becomes not operationally ready for supply reasons. No alternative loads can be defined. Reality would dictate alternative loading to accomplish a mission.

Base repair - Recovery from attack is important to sustaining a warfighting capability. A significant part of this capability is runway and taxiway repair to permit aircraft operations to be resumed after an attack on the airbase. TSAR simulates civil engineering, runway repair machinery, varying crater sizes, shelter repair, and runway repair materials as part of the overall recovery feature. SORGEN accomplishes this simulation using only a time delay, established via an algorithm, that estimates how long to cease air operations based on the number of runway and taxiway craters that must be repaired without regard to size. Personnel are indirectly addressed via a feature that permits a user specified number of parallel crater repairs to be accomplished simultaneously. No materials, machinery, or shelter repair are simulated in SORGEN.

Chemical attack - TSAR simulates a chemical warfare environment to include hospitalization and recovery of personnel, increased aircraft maintenance times because of the constraints imposed by individual protective equipment, chemical cloud dispersal, weather conditions (heat factors by season), and buddy care. SORGEN does not model any of these characteristics directly.

Cross trained personnel - Both models permit user defined cross utilization of personnel for the various maintenance tasks. TSAR is able to differentiate cross utilization via levels of proficiency and via individual tasks. SORGEN handles cross training solely as substitution of one personnel type for another without regard to task peculiarity or personnel qualifications. SORGEN permits the user to specify up to 10 personnel types for substitutions while TSAR accommodates only five.

Munition use rate controllable - TSAR permits the user to specify munition consumption, as a percentage, by mission type. SORGEN assumes a 100% consumption rate on all missions. For bombing missions, SORGEN is representative of real operations; for air-to-air missions, however, TSAR is more representative. TSAR models munition reconfiguration on landing with unused munitions returned to the stockpile while SORGEN does neither.

Speed-up procedures/task expediting - This feature is unique to TSAR. It permits a user-assigned factor to account for expediting that would occur during wartime operations. This could be accomplished in SORGEN by individually changing the task times to create a wartime unique database version. This is a feature representative of real world conditions and influences sortie production directly.

Early morning/through-flight inspections/phase inspections - TSAR permits the definition of scheduled maintenance tasks based on clock hours and flight hours. Phase inspections can be deferred and/or conducted at night. SORGEN does not permit maintenance tasks assigned by clock hours, only flight hours. There is no means of deferring a scheduled task in SORGEN or specifying what time of day it is to be accomplished. Early morning inspections and phase inspections are representative of real world operations and influence sortie production directly. The deferred phase/phase at night feature of TSAR would enhance sortie production, while lack of these features in SORGEN are less representative of the true logistics scenario.

Appendix D: SORGEN to TSAR Database Cross-Index

This appendix contains the cross reference for correlating the SORGEN database to the TSAR databases. The index is arranged according to SORGEN database\relation for easy reference to the TSAR card type and column number. The following notation is used to identify those areas where less than full equivalency exits between the two databases:

SUM - indicates that the data required by SORGEN is a summation of TSAR.

VAR - indicates that the data required by SORGEN is in various locations in the TSAR database.

MULT - indicates that the data required by SORGEN is in multiple entries in the TSAR database and may require aggregation for use in SORGEN.

NTE - indicates that there is no TSAR equivalent for the SORGEN data element.

NDE - indicates that the data is not directly equivalent, requiring translation or interpretation for use in SORGEN.

The reader is referred to the simulation model documentation for clarification of data elements where less than full equivalency exists and for a thorough explanation of the databases. The cross-index is only a tool to assist in the translation of one model's database for use by the other and is not all inclusive. Every effort is made to assure the accuracy of the data contained in these appendices. These tools should not be used alone when operating these simulation models for any purpose.

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Control/Control	"Output Comment"	CT X		NTE
Control/Control	"Simulation Lngth"	CT 1	6-10	
Control/Control	"Minimum Trials"	CT 1	11-15	
Control/Control	"Maximum Trials"	CT X		NTE
Control/Control	"% Confidence"	CT X		NTE
Control/Control	"% Max Error"	CT X		NTE
Control/Control	"Random Seed"	CT 1	21-25	
Control/Control	"Ran Gen Lock?*"	CT X		NTE
Control/Control	"Input Check ?*"	CT 2/1	11-15	NDE
Control/Control	"Input Echo ?*"	CT X		NTE
Control/Control	"Debug Level*"	CT X		NTE
Control/Control	"Output Quant*"	CT 2/1	16-20	NDE
Control/Control	"Shop Output*"	CT 2/1	16-20	NDE
Control/Control	"Res Repair Smy?"	CT 2/1	16-20	NDE
Control/Control	"Resource Output*"	CT 2/1	16-20	NDE
Control/Control	"AC Time Histgm*"	CT X		NTE
Control/Control	"Sortie Plot ?*"	CT X		NTE
Control/Control	"Aircraft Rept ?*"	CT 2/1	16-20	NDE
Control/Control	"First AC Report"	CT 2/4	VAR	NDE
Control/Control	"AC Rsched Time"	CT 2/4	VAR	NDE
Control/Control	"Last AC Report"	CT X		NTE
Control/Control	"Shop Rept ?*"	CT 2/1	16-20	NDE
Control/Control	"First Shop Rept"	CT X		NTE
Control/Control	"Shop Rsched Time"	CT X		NTE
Control/Control	"Last Shop Report"	CT X		NTE
Control/Control	"Resource Rept ?*"	CT 2/1	16-20	NDE
Control/Control	"First Res Rept"	CT X		NTE
Control/Control	"Res Rsched Time"	CT X		NTE
Control/Control	"Last Res Report"	CT X		NTE
Control/Control	"Mission Rept ?*"	CT 2/1	16-20	NDE
Control/Control	"First Misn Rept"	CT X		NTE
Control/Control	"Misn Rsched Time"	CT X		NTE
Control/Control	"Last Misn Report"	CT X		NTE
Control/Control	"Tree Stack Size"	CT X		NTE
Control/Control	"Plot Update Time"	CT X		NTE
Scenario/General	"Output Comment"	CT X		NTE
Scenario/General	"RESOURCES Ver"	CT X		NTE
Scenario/General	"RES_ORDER Ver"	CT X		NTE
Scenario/General	"TSRNA_EQUIV Ver"	CT X		NTE
Scenario/General	"Min Remain Time"	CT X		NTE
Scenario/General	"Min Accum. Time"	CT X		NTE
Scenario/General	"Priorty Intrpt?*"	CT X		NTE
Scenario/General	"Defer Tasks?*"	CT 3/1	11-15	
Scenario/General	"Auto Res Rspy?*"	CT 33	NDE	
Scenario/General	"Use X-Train?*"	CT 17/1	21-25 & 26-30	
Scenario/General	"Unsh Mnt Prb Mod"	CT 18/2		
Scenario/General	"%ShltrDmg=destyd"	CT 2/1	76-80	NDE

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Scenario/Bases	"Type*"	CT X		NTE
Scenario/Bases	"Name"	CT X		NTE
Scenario/Bases	"BASE Version"	CT X		NTE
Scenario/Bases	"ATTACK Version"	CT X		NTE
Scenario/Bases	"BASE_MODS Versn"	CT X		NTE
Scenario/Missions	"Type*"	CT 50		NDE
Scenario/Missions	"Mission Name"	CT 50	16-20	NDE
Scenario/Missions	"1st Takeoff Time"	CT 50		NDE
Scenario/Missions	"Lst Takeoff Time"	CT 50		NDE
Scenario/Missions	"Open Daily Wndow"	CT X		NTE
Scenario/Missions	"Clos Daily Wndow"	CT X		NTE
Scenario/Missions	"Resch. Interval"	CT X		NTE
Scenario/Missions	"Hours Notice"	CT 50	41-45	
Scenario/Eff_Goals	"Day"	CT X		NTE
Scenario/Eff_Goals	"Effect Goal"	CT X		NTE
Scenario/AC_Spares	"Aircraft Name"	CT 20/77	10	
Scenario/AC_Spares	"Quantity"	CT 20/77	11-15	
Scenario/AC_Spares	"Reorder Hours"	CT 20/77	16-20	UNITS NDE
Scenario/AC_Spares	"Reodr Dist Param"	CT X		NTE
Scenario/AC_Spares	"Reodr Dist Type*"	CT X		NTE
Scenario/AC_Spares	"Init Flight Hrs"	CT 41		CT 42 ALSO NDE
Scenario/AC_Spares	"Initial Status*"	CT 41		CT 42 ALSO NDE
Scenario/AC_Databases	"Aircraft Name"	CT 1	31-35	
Scenario/AC_Databases	"AC Database Name"	CT X		NTE
Base/General	"Base Type*"	CT X		NTE
Base/General	"Begin Day Shift"	CT 18/1	VAR	SHOP DEPENDENT
Base/General	"Begin Nite Shift"	CT X		NTE
Base/General	"Init POL Stocks"	CT 27	11-15	
Base/General	"POL Capacity"	CT 17/1	51-55	
Base/General	"POL Threshold"	CT X		NTE
Base/General	"POL Reord Amt"	CT 31	VAR	TYPE 0
Base/General	"Number of Rwys"	CT 17/6	11-15	NDE
Base/General	"Number of Nodes"	CT 17/3	11-15	
Base/General	"Number of Arcs"	CT 17/3	16-20	
Base/General	"Number of Ramps"	CT 17/3	21-25	
Base/General	"Number Shelters"	CT 17/1	36-40	
Base/General	"Pre Taxi Time"	CT 17/1	66-70	
Base/General	"Post Taxi Time"	CT 15/1	11-14	
Base/General	"Survey/EOD Time"	CT 17/9	26-30	
Base/General	"PA Task Delay"	CT 17/9	16-25	
Base/General	"Crater Rep Time"	CT 38	11-13	NDE
Base/General	"Dist Parameter"	CT X		NDE
Base/General	"Dist Type*"	CT X		NDE

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Base/General	"#Parallel Reprs"	CT X		NDE
Base/General	"MCL"	CT 17/7	31-35	
Base/General	"MCW"	CT 17/7	36-40	
Base/General	"Extended MCL"	CT 17/7	16-20	
Base/General	"Extended MCW"	CT 17/7	21-25	
Base/General	"Max Runways"	CT 17/7	26-30	
Base/General	"RRMODE"	CT 17/7	11-15	
Base/Taxiarc	"1st Node"	CT 17/4	VAR	MULTI DATA ON CARD
Base/Taxiarc	"2nd Node"	CT 17/4	VAR	MULTI DATA ON CARD
Base/Taxiarc	"Length"	CT 17/4	VAR	MULTI DATA ON CARD
Base/Runways	"Runway #"	CT 17/6	11-15	
Base/Runways	"1st ARC #"	CT 17/6	16-20	FIRST CARD IMAGE
Base/Runways	"2nd ARC #"	CT 17/6	21-25	FIRST CARD IMAGE
Base/Runways	"3rd ARC #"	CT 17/6	26-30	FIRST CARD IMAGE
Base/Runways	"4th ARC #"	CT 17/6	31-35	FIRST CARD IMAGE
Base/Runways	"5th ARC #"	CT 17/6	36-40	FIRST CARD IMAGE
Base/Runways	"6th ARC #"	CT 17/6	41-45	FIRST CARD IMAGE
Base/Runways	"7th ARC #"	CT 17/6	46-50	FIRST CARD IMAGE
Base/Runways	"8th ARC #"	CT 17/6	51-55	FIRST CARD IMAGE
Base/Runways	"9th ARC #"	CT 17/6	56-60	FIRST CARD IMAGE
Base/Runways	"10th ARC #"	CT 17/6	61-65	FIRST CARD IMAGE
Base/Runways	"11th ARC #"	CT 17/6	16-20	SECOND CARD IMAGE
Base/Runways	"12th ARC #"	CT 17/6	21-25	SECOND CARD IMAGE
Base/Runways	"13th ARC #"	CT 17/6	26-30	SECOND CARD IMAGE
Base/Runways	"14th ARC #"	CT 17/6	31-35	SECOND CARD IMAGE
Base/Runways	"15th ARC #"	CT 17/6	36-40	SECOND CARD IMAGE
Base/Runways	"16th ARC #"	CT 17/6	41-45	SECOND CARD IMAGE
Base/Runways	"17th ARC #"	CT 17/6	46-50	SECOND CARD IMAGE
Base/Runways	"18th ARC #"	CT 17/6	51-55	SECOND CARD IMAGE
Base/Runways	"19th ARC #"	CT 17/6	56-60	SECOND CARD IMAGE
Base/Runways	"20th ARC #"	CT 17/6	61-65	SECOND CARD IMAGE
Base/Runways	"21th ARC #"	CT 17/6	16-20	THIRD CARD IMAGE
Base/Runways	"22th ARC #"	CT 17/6	21-25	THIRD CARD IMAGE
Base/Runways	"23th ARC #"	CT 17/6	26-30	THIRD CARD IMAGE
Base/Runways	"24th ARC #"	CT 17/6	31-35	THIRD CARD IMAGE
Base/Runways	"25th ARC #"	CT 17/6	36-40	THIRD CARD IMAGE
Base/Runways	"26th ARC #"	CT 17/6	41-45	THIRD CARD IMAGE
Base/Runways	"27th ARC #"	CT 17/6	46-50	THIRD CARD IMAGE
Base/Runways	"28th ARC #"	CT 17/6	51-55	THIRD CARD IMAGE
Base/Runways	"29th ARC #"	CT 17/6	56-60	THIRD CARD IMAGE
Base/Runways	"30th ARC #"	CT 17/6	61-65	THIRD CARD IMAGE
Base/Runways	"31th ARC #"	CT 17/6	16-20	THIRD CARD IMAGE
Base/Runways	"32th ARC #"	CT 17/6	21-25	THIRD CARD IMAGE
Base/Runways	"33th ARC #"	CT 17/6	26-30	THIRD CARD IMAGE
Base/Runways	"34th ARC #"	CT 17/6	31-35	THIRD CARD IMAGE
Base/Runways	"35th ARC #"	CT 17/6	35-40	THIRD CARD IMAGE
Base/Runways	"36th ARC #"	CT 17/6	41-45	THIRD CARD IMAGE
Base/Runways	"37th ARC #"	CT 17/6	46-50	THIRD CARD IMAGE

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Base/Shelters	"Nearest Node"	CT 17/5	VAR	CC 10-15 TO 61-65
Base/Ramps	"Relative Cap."	CT 17/8	21-25	CC 15 = 1
Base/Ramps				
	"Nearest Node"	CT 17/8	21-25	CC 15 = 2
Base/AC_Basing	"Aircraft Name"	CT 20	6-9	
Base/AC_Basing	"Quantity"	CT 20	11-15	
Base/AC_Basing	"Arrival Time"	CT 20/66	6-10	NOTE UNITS
Base/AC_Basing	"Init Flight Hrs"	CT X		NTE
Base/AC_Basing	"Initial Status*"	CT 41		NDE
Base/AC_Basing	"Mission Config"	CT 41		NDE
Base/AC_Networks	"Aircraft Name"	CT X		NTE
Base/AC_Networks	"Task Field 1"	CT 29	16-20	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 2"	CT 29	21-25	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 3"	CT 29	26-30	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 4"	CT 29	31-35	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 5"	CT 29	36-40	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 6"	CT 29	41-45	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 7"	CT 29	46-50	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 8"	CT 29	51-55	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 9"	CT 29	56-60	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 10"	CT 29	61-65	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 11"	CT 29	16-20	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 12"	CT 29	21-25	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 13"	CT 29	26-30	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 14"	CT 29	31-35	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 15"	CT 29	36-40	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 16"	CT 29	41-45	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 17"	CT 29	46-50	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 18"	CT 29	51-55	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 19"	CT 29	56-60	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 20"	CT 29	61-65	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 21"	CT 29	16-20	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 22"	CT 29	21-25	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 23"	CT 29	26-30	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 24"	CT 29	31-35	THIRD CAK IMAGE
Base/AC_Networks	"Task Field 25"	CT 29	36-40	THIRD CARD iMAGE
Base/AC_Networks	"Task Field 26"	CT 29	41-45	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 27"	CT 29	46-50	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 28"	CT 29	51-55	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 29"	CT 29	56-60	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 30"	CT 29	61-65	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 31"	CT 29	16-20	FOURTH CARD IMAGE
Base/AC_Networks	"Task Field 32"	CT 29	21-25	FOURTH CARD IMAGE
Base/AC_Networks	"Task Field 33"	CT 29	26-30	FOURTH CARD IMAGE
Base/AC_Networks	"Task Field 34"	CT 29	31-35	FOURTH CARD IMAGE
Base/AC_Networks	"Task Field 35"	CT 29	36-40	FOURTH CARD IMAGE
Base/AC_Networks	"Task Field 36"	CT 29	41-45	FOURTH CARD IMAGE

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Base/AC_Networks	"Task Field 37"	CT 29	46-50	FOURTH CARD IMAGE
Base/Personnel	"Personnel Name"	CT 21	VAR	NDE
Base/Personnel	"Initial Number"	CT 21	VAR	NDE
Base/Personnel	"Target Number"	CT 21	VAR	NDE
Base/Personnel	"Min Crew Size"	CT 21	VAR	NDE
Base/Personnel	"% Day Shift"	CT 21	VAR	NDE
Base/Cross_Train	"Requested Type"	CT 45/2	6-10	
Base/Cross_Train	"CT Type # 1"	CT 45/2	11-15	
Base/Cross_Train	"CT Type # 2"	CT 45/2	16-20	
Base/Cross_Train	"CT Type # 3"	CT 45/2	21-25	
Base/Cross_Train	"CT Type # 4"	CT 45/2	26-30	
Base/Cross_Train	"CT Type # 5"	CT 45/2	31-35	
Base/Cross_Train	"CT Type # 6"	CT X	TSAR LIMITED TO 5	
Base/Cross_Train	"CT Type # 7"	CT X	TSAR LIMITED TO 5	
Base/Cross_Train	"CT Type # 8"	CT X	TSAR LIMITED TO 5	
Base/Cross_Train	"CT Type # 9"	CT X	TSAR LIMITED TO 5	
Base/Cross_Train	"CT Type #10"	CT X	TSAR LIMITED TO 5	
Base/Parts	"Part Name"	CT 23	6-10	
Base/Parts	"Initial Number"	CT 23	12-13	
Base/Parts	"Min Inventory"	CT 23	21-25	
Base/Parts	"Reordr Threshold"	CT X		NTE
Base/Parts	"Reorder Quantity"	CT 31	VAR	26-30 THRU 71-75
Base/AGE	"AGE Name"	CT 22	6-10	
Base/AGE	"Initial Number"	CT 22	12-13	
Base/AGE	"Min Inventory"	CT X		NTE
Base/AGE	"Reordr Threshold"	CT X		NTE
Base/AGE	"Reorder Quantity	CT X	VAR	26-30 THRU 71-75
Base/TRAP	"TRAP Name"	CT 25	VAR	6-10 THRU 46-50
Base/TRAP	"Initial Number"	CT 25	VAR	11-15 THRU 51-55
Base/TRAP	"Min Inventory"	CT X		NTE
Base/TRAP	"Reordr Threshold"	CT X		NTE
Base/TRAP	"Reorder Quantity"	CT 31	VAR	26-30 THRU 71-75
Base/Munitions	"Munition Name"	CT 24	VAR	6-10 THRU 66-70
Base/Munitions	"Initial Number"	CT 24	VAR	11-15 THRU 71-75
Base/Munitions	"Min Inventory"	CT X		NTE
Base/Munitions	"Reordr Threshold"	CT X		NTE
Base/Munitions	"Reorder Quantity	CT 31	VAR	26-30 THRU 71-75
Mission/Mission	"Priority"	CT 50	21-25	
Mission/Mission	"Aircraft Type"	CT 50	11-15	
Mission/Mission	"Desired # AC"	CT 50	31-35	
Mission/Mission	"Minimum # AC"	CT 50	36-40	
Mission/Mission	"Avg Config Time"	SUM	VAR	CT 13,14,15

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Mission/Mission	"Config Dist Par"	CT X		NTE
Mission/Mission	"Config Dist Typ*"	CT X		NTE
Mission/Mission	"Shop"	CT 13	31-33	SUM
		CT 14	26-28	SUM
Mission/Mission	"Pers #1 Type"	CT 13	31-33	SUM
		CT 14	26-28	SUM
Mission/Mission	"Pers #1 Quantity"	CT 13	34-35	SUM
		CT 14	29-30	SUM
Mission/Mission	"Pers #2 Type"	CT 13	61-63	SUM
		CT 14	56-58	SUM
Mission/Mission	"Pers #2 Quantity"	CT 13	64-65	SUM
		CT 14	59-60	SUM
Mission/Mission	"AGE Type"	CT 13 & 14	VAR	SUM
Mission/Mission	"AGE Quantity"	CT 13 & 14	VAR	SUM
Mission/Mission	"Primary Munition"	CT 13	16-18	
Mission/Mission	"PM Quantity"	CT 13	19-20	
Mission/Mission	"Second Munition"	CT 13	46-48	
Mission/Mission	"SM Quantity"	CT 13	49-50	
Mission/Mission	"TRAP Type"	CT 14	11-14 & 41-44	
Mission/Mission	"TRAP Quantity"	CT 14	15 & 45	
Mission/Mission	"PM Effect Val"	CT X		NTE
Mission/Mission	"PM Effect Val"	CT X		NTE
Mission/Mission	"Mission Window"	CT X		NTE
Mission/Mission	"Prob Gnd Abort"	CT 16	31-35	NDE
Mission/Mission	"Pkill"	CT 16	23-26	NDE
Mission/Mission	"Prd"	CT 16	19-22	NDE
Mission/Mission	"Pnrd"	CT 16	19-22	NDE
Mission/Mission	"Avg Sortie Time"	CT 16	11-14	
Mission/Mission	"Sortie Dist Par"	CT 16	15	NDE
Mission/Mission	"Sortie Dist Typ*"	CT 16	15	NDE
Mission/Attr_Modifiers	"Function Type*"	CT 16	VAR	NDE
Mission/Attr_Modifiers	"Day of Scenario"	CT 16	VAR	51-52 THRU 71-72
Mission/Attr_Modifiers	"Effect. Value"	CT X		NTE
Mission/Attr_Modifiers	"Pkill_Mul"	CT 16	51-75	NDE
Mission/Attr_Modifiers	"Prd_Mul"	CT 16	51-75	NDE
Mission/Attr_Modifiers	"Pnrd_Mul"	CT 16	51-75	NDE
Aircraft/General	"Max Defer Tasks"	CT 3/1	26-30	NDE
Aircraft/TSKRQT	"Name"	CT 7	3-7	NDE
Aircraft/TSKRQT	"Shop"	CT 5	8-9	NDE
Aircraft/TSKRQT	"Root Task ?*"	CT 5	37-40	YES IF VALUE > 0
Aircraft/TSKRQT	"Failure Mech*"	CT X		TSAR IS ALL PROB
Aircraft/TSKRQT	"Fail Mech Value"	MULT	VAR	CT 7 ROOT CT 5 NET
Aircraft/TSKRQT	"Deferability"	CT 3/1	26-30	
Aircraft/TSKRQT	"MTTR"	CT 5	16-19	
Aircraft/TSKRQT	"Dist Parameter"	CT 5	20	NDE
Aircraft/TSKRQT	"Dist Type*"	CT 5	20	NDE

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Aircraft/TSKRQT	"Task Location*"	CT 5	20	NDE
Aircraft/TSKRQT	"LRU/Consumable"	CT 5	11-15	
Aircraft/TSKRQT	"LC Quantity"	CT X		NTE (ASSUME 1)
Aircraft/TSKRQT	"LC Probability"	CT 5	49-51	
Aircraft/TSKRQT	"Pers #1 Type"	CT 5	21-23	
Aircraft/TSKRQT	"Pers #1 Quantity"	CT 5	24-25	
Aircraft/TSKRQT	"Pers #2 Type"	CT 5	26-28	
Aircraft/TSKRQT	"Pers #2 Quantity"	CT 5	29-30	
Aircraft/TSKRQT	"AGE Type"	CT 5	31-33	NDE
Aircraft/TSKRQT	"AGE Quantity"	CT X		NDE ASSUME 1
Aircraft/TSKRQT	"Unscheduled*"	CT 5	45	NDE
Aircraft/Network	"Base Task"	CT 5	37-40	YES IF > 0 NDE
Aircraft/Network	"Spawned Task"	CT 5	41-44	YES IF <> 0NDE
Aircraft/Network	"Mutually Exc*"	CT 5	41-44	YES WHEN NEGATIVE
Aircraft/Network	"Prob Spawn"	CT 5	41-44	NDE (÷ BY 1000)
Resource/Resources	"Resource Name"	CT 8/1	6-10	NDE
Resource/Resources	"Priority"	CT X		NTE
Resource/Resources	"Resource Type*"	CT X		NTE
Resource/Resources	"Parent LRU"	CT 8/2	21-25	
Resource/Resources	"Failure Mech*"	CT X		NTE (SET TO PROB)
Resource/Resources	"Fail Mech Value"	CT 5	49-51	NDE (÷ BY 100)
Resource/Resources	"Quant per LRU"	CT X		NTE (ASSUME 1)
Resource/Resources	"Loc[illegible] Repair*"	CT X		NTE
Resource/Resources	"Base Repair Time"	CT 8/1	16-19	
Resource/Resources	"Base Dist Param"	CT 8/1	20	
Resource/Resources	"Base Dist Type*"	CT 8/1	20	
Resource/Resources	"Base Pers Type"	CT 8/1	21-23	
Resource/Resources	"Base Pers Quant"	CT 8/1	24-25	
Resource/Resources	"Base AGE Type"	CT 8/1	26-30	
Resource/Resources	"Base Condemned"	CT 8/1	56-60	
Resource/Resources	"Base NRTS Rate"	CT 23/2XX	VAR	26-30 THRU 76-80
Resource/Resources	"CIRF Repair Time"	CT 8/3	16-19	
Resource/Resources	"CIRF Dist Param"	CT 8/3	20	NDE
Resource/Resources	"CIRF Dist Type*"	CT 8/3	20	NDE
Resource/Resources	"CIRF Pers Type"	CT 8/3	21-23	
Resource/Resources	"CIRF Pers Quant"	CT 8/3	24-25	
Resource/Resources	"CIRF AGE Type"	CT 8/3	26-30	
Resource/Resources	"CIRF Condemned"	CT X		NTE
Resource/Resources	"CIRF NRTS Rate"	CT X		NTE
Resource/Resources	"Depo Repair Time"	CT X		NTE
Resource/Resources	"Depo Dist Param "	CT X		NTE
Resource/Resources	"Depo Dist Type*"	CT X		NTE
Resource/Resources	"Depo Pers Type"	CT X		NTE
Resource/Resources	"Depo Pers Quant"	CT X		NTE
Resource/Resources	"Depo AGE Type"	CT X		NTE
Resource/Resources	"Depo Condemned"	CT X		NTE
Resource/Resources	"Resupply Time"	CT X		NTE

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Resource/Resources	"Cost"	CT X		NTE
Resource/Resources	"WUC"	CT X		NTE
Resource/Resources	"Shop"	CT 8/1	11-15	
Resource/Resources	"Pallet Equiv"	CT X		NTE
Res_Order/Rtype_Order	"AGE Arriv Time"	CT 33	VAR	11-15 TO 71-75 NDE
Res_Order/Rtype_Order	"AGE Dist Param"	CT 33	VAR	16-20 TO 76-80 NDE
Res_Order/Rtype_Order	"AGE Dist Type*"	CT 33	VAR	16-20 TO 76-80 NDE
Res_Order/Rtype_Order	"PART Arriv Time"	CT 33	VAR	11-15 TO 71-75 NDE
Res_Order/Rtype_Order	"PART Dist Param"	CT 33		
Res_Order/Rtype_Order	"PART Dist Type*"	CT 33		
Res_Order/Rtype_Order	"MUN Arriv Time"	CT 31	6-15	NDE
Res_Order/Rtype_Order	"MUN Dist Param"	CT X		NTE
Res_Order/Rtype_Order	"MUN Dist Type*"	CT X		NTE
Res_Order/Rtype_Order	"TRAP Arriv Time"	CT 33	VAR	11-15 TO 71-75 NDE
Res_Order/Rtype_Order	"TRAP Dist Param"	CT 33	VAR	16-20 TO 76-80 NDE
Res_Order/Rtype_Order	"TRAP Dist Type*"	CT 33	VAR	16-20 TO 76-80 NDE
Res_Order/Rtype_Order	"POL Arriv Time"	CT 31	6-15	NDE
Res_Order/Rtype_Order	"POL Dist Param"	CT X		NTE
Res_Order/Rtype_Order	"POL Dist Type*"	CT X		NTE
Res_Order/Res_Order	"Resource Name"	CT 31	VAR	21-25 THRU 66-70
Res_Order/Res_Order	"Resource Type*"	CT 31	VAR	31-35 THRU 76-80
Res_Order/Res_Order	"Day"	CT 31	6-10	
Res_Order/Res_Order	"Time-to-Arrive"	CT 31	11-15	
Res_Order/Res_Order	"Dist Param"	CT X		NTE
Res_Order/Res_Order	"Dist Type*"	CT X		NTE
Base_Mods/Base_Mods	"Time"	CT X		NTE
Base_Mods/Base_Mods	"Resource Name"	CT X		NTE
Base_Mods/Base_Mods	"Resource Type*"	CT X		NTE
Base_Mods/Base_Mods	"Change Type*"	CT X		NTE
Base_Mods/Base_Mods	"Magnitude"	CT X		NTE

Appendix E: TSAR to SORGEN Database Cross-Index

This appendix contains the cross reference for correlating the TSAR database to the SORGEN databases. The index is arranged according to TSAR card type and column number for easy reference back to the SORGEN data element. The following notation is used to identify those areas where less than full equivalency exits between the two databases:

SUM - indicates that the data required by SORGEN is a summation of TSAR data which may be multiple locations.

VAR - indicates that the data required by SORGEN is in various locations in the TSAR database.

MULT - indicates that the data required by SORGEN is in multiple entries in the TSAR database and may requrie aggregation for use in SORGEN.

NTE - indicates that there is no TSAR equivalent for the SORGEN data element.

NDE - indicates that the data is not directly equivalent, requiring translation or interpretation for use in SORGEN.

The reader is referred to the simulation model documentation for clarification of data elements where less than full equivalency exists and for a thorough explanation of the databases. The cross-index is only a tool to assist in the translation of one model's database for use by the other and is not all inclusive. Every effort is made to assure the accuracy of the data contained in these appendices. These tools should not be used alone when operating these simulation models for any purpose.

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Base/Parts	"Reordr Threshold"	CT X		NTE
Scenario/Missions	"Open Daily Wndow"	CT X		NTE
Base/AGE	"Reordr Threshold"	CT X		NTE
Scenario/Bases	"BASE_MODS Versn"	CT X		NTE
Scenario/Missions	"Clos Daily Wndow"	CT X		NTE
Scenario/Bases	"Name"	CT X		NTE
Scenario/Bases	"ATTACK Version"	CT X		NTE
Scenario/Bases	"Type*"	CT X		NTE
Base/AGE	"Min Inventory"	CT X		NTE
Scenario/AC_Spares	"Reodr Dist Param"	CT X		NTE
Base/AC_Basing	"Init Flight Hrs"	CT X		NTE
Base/AC_Networks	"Aircraft Name"	CT X		NTE
Base/General	"Dist Parameter"	CT X		NTE
Base/General	"Dist Type*"	CT X		NTE
Control/Control	"Output Comment"	CT X		NTE
Base/General	"#Parallel Reprs"	CT X		NTE
Base/General	"POL Threshold"	CT X		NTE
Base/General	"Begin Nite Shift"	CT X		NTE
Scenario/Eff_Goals	"Effect Goal"	CT X		NTE
Scenario/Eff_Goals	"Day"	CT X		NTE
Scenario/General	"Priorty Intrpt?*"	CT X		NTE
Scenario/AC_Spares	"Reodr Dist Type*"	CT X		NTE
Base/General	"Base Type*"	CT X		NTE
Scenario/AC_Databases	"AC Database Name"	CT X		NTE
Scenario/Missions	"Resch. Interval"	CT X		NTE
Scenario/Bases	"BASE Version"	CT X		NTE
Control/Control	"First Res Rept"	CT X		NTE
Control/Control	"Ran Gen Lock?*"	CT X		NTE
Control/Control	"Res Rsched Time"	CT X		NTE
Control/Control	"Last Res Report"	CT X		NTE
Control/Control	"Sortie Plot ?*"	CT X		NTE
Control/Control	"Last Shop Report"	CT X		NTE
Control/Control	"Shop Rsched Time"	CT X		NTE
Control/Control	"Debug Level*"	CT X		NTE
Control/Control	"AC Time Histgm*"	CT X		NTE
Control/Control	"Last AC Report"	CT X		NTE
Control/Control	"Input Echo ?*"	CT X		NTE
Control/Control	"First Shop Rept"	CT X		NTE
Control/Control	"First Misn Rept"	CT X		NTE
Control/Control	"% Max Error"	CT X		NTE
Control/Control	"Misn Rsched Time"	CT X		NTE
Scenario/General	"TSRNA_EQUIV Ver"	CT X		NTE
Control/Control	"% Confidence"	CT X		NTE
Control/Control	"Maximum Trials"	CT X		NTE
Scenario/General	"Min Accum. Time"	CT X		NTE
Scenario/General	"RES_ORDER Ver"	CT X		NTE
Scenario/General	"Min Remain Time"	CT X		NTE
Scenario/General	"RESOURCES Ver"	CT X		NTE
Control/Control	"Tree Stack Size"	CT X		NTE

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Control/Control	"Last Misn Report"	CT X		NTE
Scenario/General	"Output Comment"	CT X		NTE
Control/Control	"Plot Update Time"	CT X		NTE
Resource/Resources	"Depo Dist Type*"	CT X		NTE
Resource/Resources	"Priority"	CT X		NTE
Resource/Resources	"Depo Dist Param "	CT X		NTE
Aircraft/TSKRQT	"AGE Quantity"	CT X		NDE ASSUME 1
Aircraft/TSKRQT	"Failure Mech*"	CT X		TSAR IS ALL PROB
Resource/Resources	"CIRF NRTS Rate"	CT X		NTE
Aircraft/TSKRQT	"LC Quantity"	CT X		NTE (ASSUME 1)
Resource/Resources	"Depo Repair Time"	CT X		NTE
Resource/Resources	"Depo Pers Type"	CT X		NTE
Resource/Resources	"Resource Type*"	CT X		NTE
Resource/Resources	"Cost"	CT X		NTE
Resource/Resources	"Resupply Time"	CT X		NTE
Resource/Resources	"Depo Condemned"	CT X		NTE
Resource/Resources	"Depo AGE Type"	CT X		NTE
Resource/Resources	"WUC"	CT X		NTE
Resource/Resources	"Locat of Repair*"	CT X		NTE
Resource/Resources	"Pallet Equiv"	CT X		NTE
Resource/Resources	"Failure Mech*"	CT X		NTE (SET TO PROB)
Resource/Resources	"Quant per LRU"	CT X		NTE (ASSUME 1)
Resource/Resources	"Depo Pers Quant"	CT X		NTE
Resource/Resources	"CIRF Condemned"	CT X		NTE
Res_Order/Rtype_Order	"POL Dist Param"	CT X		NTE
Mission/Attr_Modifiers	"Effect. Value"	CT X		NTE
Res_Order/Rtype_Order	"POL Dist Type*"	CT X		NTE
Mission/Mission	"Config Dist Typ*"	CT X		NTE
Mission/Mission	"Config Dist Par"	CT X		NTE
Base/Munitions	"Min Inventory"	CT X		NTE
Base/Munitions	"Reordr Threshold"	CT X		NTE
Res_Order/Res_Order	"Dist Param"	CT X		NTE
Res_Order/Res_Order	"Dist Type*"	CT X		NTE
Base_Mods/Base_Mods	"Change Type*"	CT X		NTE
Mission/Mission	"FM Effect Val"	CT X		NTE
Mission/Mission	"Mission Window"	CT X		NTE
Base_Mods/Base_Mods	"Resource Type*"	CT X		NTE
Base_Mods/Base_Mods	"Resource Name"	CT X		NTE
Base_Mods/Base_Mods	"Time"	CT X		NTE
Base_Mods/Base_Mods	"Magnitude"	CT X		NTE
Res_Order/Rtype_Order	"MUN Dist Type*"	CT X		NTE
Mission/Mission	"PM Effect Val"	CT X		NTE
Base/TRAP	"Min Inventory"	CT X		NTE
Res_Order/Rtype_Order	"MUN Dist Param"	CT X		NTE
Base/TRAP	"Reordr Threshold"	CT X		NTE
Base/Cross_Train	"CT Type # 6"	CT X	TSAR LIMITED TO 5	
Base/Cross_Train	"CT Type # 7"	CT X	TSAR LIMITED TO 5	
Base/Cross_Train	"CT Type # 8"	CT X	TSAR LIMITED TO 5	

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Base/Cross_Train	"CT Type # 9"	CT X	TSAR LIMITED TO 5	
Base/Cross_Train	"CT Type # 10"	CT X	TSAR LIMITED TO 5	
Control/Control	"Simulation Lngth"	CT 1	6-10	
Control/Control	"Minimum Trials"	CT 1	11-15	
Control/Control	"Random Seed"	CT 1	21-25	
Scenario/AC_Databases	"Aircraft Name"	CT 1	31-35	
Control/Control	"Input Check ?*"	CT 2 /1	11-15	NDE
Control/Control	"Mission Rept ?*"	CT 2 /1	16-20	NDE
Control/Control	"Resource Rept ?*"	CT 2 /1	16-20	NDE
Control/Control	"Shop Rept ?*"	CT 2 /1	16-20	NDE
Control/Control	"Output Quant*"	CT 2 /1	16-20	NDE
Control/Control	"Shop Output*"	CT 2 /1	16-20	NDE
Control/Control	"Resource Output*"	CT 2 /1	16-20	NDE
Control/Control	"Aircraft Rept ?*"	CT 2 /1	16-20	NDE
Control/Control	"Res Repair Smy?"	CT 2 /1	16-20	NDE
Scenario/General	"%ShltrDmg=destyd"	CT 2 /1	76-80	NDE
Control/Control	"First AC Report"	CT 2 /4	VAR	NDE
Control/Control	"AC Rsched Time"	CT 2 /4	VAR	NDE
Scenario/General	"Defer Tasks?*"	CT 3 /1	11-15	
Aircraft/TSKRQT	"Deferability"	CT 3 /1	26-30	
Aircraft/General	"Max Defer Tasks"	CT 3 /1	26-30	NDE
Aircraft/TSKRQT	"Fail Mech Value"	CT 5	41-45	NETWORK TASKS
Aircraft/TSKRQT	"Task Location*"	CT 5	20	NDE
Aircraft/TSKRQT	"Dist Type*"	CT 5	20	NDE
Aircraft/TSKRQT	"Dist Parameter"	CT 5	20	NDE
Aircraft/TSKRQT	"Unscheduled*"	CT 5	45	NDE
Aircraft/TSKRQT	"Shop"	CT 5	8-9	NDE
Aircraft/TSKRQT	"LRU/Consumable"	CT 5	11-15	
Aircraft/TSKRQT	"MTTR"	CT 5	16-19	
Aircraft/TSKRQT	"Pers #1 Type"	CT 5	21-23	
Aircraft/TSKRQT	"Pers #1 Quantity"	CT 5	24-25	
Aircraft/TSKRQT	"Pers #2 Type"	CT 5	26-28	
Aircraft/TSKRQT	"Pers #2 Quantity"	CT 5	29-30	
Aircraft/TSKRQT	"AGE Type"	CT 5	31-33	NDE
Aircraft/TSKRQT	"Root Task ?*"	CT 5	37-40	YES IF VALUE > 0
Aircraft/Network	"Base Task"	CT 5	37-40	YES IF > 0 NDE
Aircraft/Network	"Mutually Exc*"	CT 5	41-44	YES WHEN NEGATIVE
Aircraft/Network	"Prob Spawn"	CT 5	41-44	NDE (÷ BY 1000)
Aircraft/Network	"Spawned Task"	CT 5	41-44	YES IF <> ONDE
Aircraft/TSKRQT	"LC Probability"	CT 5	49-51	
Resource/Resources	"Fail Mech Value"	CT 5	49-51	NDE (÷ BY 100)
Aircraft/TSKRQT	"Fail Mech Value"	CT 7	VAR	ROOT TASKS
Aircraft/TSKRQT	"Name"	CT 7	3-7	NDE
Resource/Resources	"Base Dist Type*"	CT 8 /1	20	
Resource/Resources	"Base Dist Param"	CT 8 /1	20	
Resource/Resources	"Resource Name"	CT 8 /1	6-10	NDE
Resource/Resources	"Shop"	CT 8 /1	11-15	
Resource/Resources	"Base Repair Time"	CT 8 /1	16-19	
Resource/Resources	"Base Pers Type"	CT 8 /1	21-23	

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Resource/Resources	"Base Pers Quant"	CT 8 /1	24-25	
Resource/Resources	"Base AGE Type"	CT 8 /1	26-30	
Resource/Resources	"Base Condemned"	CT 8 /1	56-60	
Resource/Resources	"Parent LRU"	CT 8 /2	21-25	
Resource/Resources	"CIRF Dist Type*"	CT 8 /3	20	MDE
Resource/Resources	"CIRF Dist Param"	CT 8 /3	20	MDE
Resource/Resources	"CIRF Repair Time"	CT 8 /3	16-19	
Resource/Resources	"CIRF Pers Type"	CT 8 /3	21-23	
Resource/Resources	"CIRF Pers Quant"	CT 8 /3	24-25	
Resource/Resources	"CIRF AGE Type"	CT 8 /3	26-30	
Mission/Mission	"Primary Munition"	CT 13	16-18	
Mission/Mission	"PM Quantity"	CT 13	19-20	
Mission/Mission	"Second Munition"	CT 13	46-48	
Mission/Mission	"SM Quantity"	CT 13	49-50	
Mission/Mission	"Pers #2 Type"	CT 13 61-63	VAR	SUM
		CT 14 56-58	VAR	SUM
Mission/Mission	"Pers #1 Quantity"	CT 13 34-35	VAR	SUM
		CT 14 29-30	VAR	SUM
Mission/Mission	"Pers #2 Quantity"	CT 13 64-65	VAR	SUM
		CT 14 59-60	VAR	SUM
Mission/Mission	"Shop"	CT 13 31-33	VAR	SUM
		CT 14 26-28	VAR	SUM
Mission/Mission	"Pers #1 Type"	CT 13 31-33	VAR	SUM
		CT 14 26-28	VAR	SUM
Mission/Mission	"AGE Type"	CT 13 & 14	VAR	SUM
Mission/Mission	"Avg Config Time"	CT 13,14,15	VAR	SUM
Mission/Mission	"AGE Quantity"	CT 13 & 14	VAR	SUM
Mission/Mission	"TRAP Quantity"	CT 14	15 & 45	
Mission/Mission	"TRAP Type"	CT 14	11-14 & 41-44	
Base/General	"Post Taxi Time"	CT 15 /1	11-14	
Mission/Attr_Modifiers	"Function Type*"	CT 16	VAR	MDE
Mission/Attr_Modifiers	"Day of Scenario"	CT 16	VAR	51-52 THRU 71-72
Mission/Mission	"Sortie Dist Par"	CT 16	15	MDE
Mission/Mission	"Sortie Dist Typ*"	CT 16	15	MDE
Mission/Mission	"Avg Sortie Time"	CT 16	11-14	
Mission/Mission	"Pnrd"	CT 16	19-22	MDE
Mission/Mission	"Prd"	CT 16	19-22	MDE
Mission/Mission	"Pkill"	CT 16	23-26	MDE
Mission/Mission	"Prob Gnd Abort"	CT 16	31-35	MDE
Mission/Attr_Modifiers	"Pnrd_Mul"	CT 16	51-75	MDE
Mission/Attr_Modifiers	"Pkill_Mul"	CT 16	51-75	MDE
Mission/Attr_Modifiers	"Prd_Mul"	CT 16	51-75	MDE
Scenario/General	"Use X-Train?*"	CT 17 /1	21-25 & 26-30	
Base/General	"Number Shelters"	CT 17 /1	36-40	
Base/General	"POL Capacity"	CT 17 /1	51-55	
Base/General	"Pre Taxi Time"	CT 17 /1	66-70	
Base/General	"Number of Nodes"	CT 17 /3	11-15	
Base/General	"Number of Arcs"	CT 17 /3	16-20	
Base/General	"Number of Ramps"	CT 17 /3	21-25	

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Base/Taxiarc	"1st Node"	CT 17 /4	VAR	MULTI DATA ON CARD
Base/Taxiarc	"2nd Node"	CT 17 /4	VAR	MULTI DATA ON CARD
Base/Taxiarc	"Length"	CT 17 /4	VAR	MULTI DATA ON CARD
Base/Shelters	"Nearest Node"	CT 17 /5	VAR	10-15 TO 61-65
Base/Runways	"Runway #"	CT 17 /6	11-15	
Base/General	"Number of Rwys"	CT 17 /6	11-15	NODE
Base/Runways	"1st ARC #"	CT 17 /6	16-20	FIRST CARD IMAGE
Base/Runways	"2nd ARC #"	CT 17 /6	21-25	FIRST CARD IMAGE
Base/Runways	"3rd ARC #"	CT 17 /6	26-30	FIRST CARD IMAGE
Base/Runways	"4th ARC #"	CT 17 /6	31-35	FIRST CARD IMAGE
Base/Runways	"5th ARC #"	CT 17 /6	36-40	FIRST CARD IMAGE
Base/Runways	"6th ARC #"	CT 17 /6	41-45	FIRST CARD IMAGE
Base/Runways	"7th ARC #"	CT 17 /6	46-50	FIRST CARD IMAGE
Base/Runways	"8th ARC #"	CT 17 /6	51-55	FIRST CARD IMAGE
Base/Runways	"9th ARC #"	CT 17 /6	56-60	FIRST CARD IMAGE
Base/Runways	"10th ARC #"	CT 17 /6	61-65	FIRST CARD IMAGE
Base/Runways	"11th ARC #"	CT 17 /6	16-20	SECOND CARD IMAGE
Base/Runways	"12th ARC #"	CT 17 /6	21-25	SECOND CARD IMAGE
Base/Runways	"13th ARC #"	CT 17 /6	26-30	SECOND CARD IMAGE
Base/Runways	"14th ARC #"	CT 17 /6	31-35	SECOND CARD IMAGE
Base/Runways	"15th ARC #"	CT 17 /6	36-40	SECOND CARD IMAGE
Base/Runways	"16th ARC #"	CT 17 /6	41-45	SECOND CARD IMAGE
Base/Runways	"17th ARC #"	CT 17 /6	46-50	SECOND CARD IMAGE
Base/Runways	"18th ARC #"	CT 17 /6	51-55	SECOND CARD IMAGE
Base/Runways	"19th ARC #"	CT 17 /6	56-60	SECOND CARD IMAGE
Base/Runways	"20th ARC #"	CT 17 /6	61-65	SECOND CARD IMAGE
Base/Runways	"21th ARC #"	CT 17 /6	16-20	THIRD CARD IMAGE
Base/Runways	"22th ARC #"	CT 17 /6	21-25	THIRD CARD IMAGE
Base/Runways	"23th ARC #"	CT 17 /6	26-30	THIRD CARD IMAGE
Base/Runways	"24th ARC #"	CT 17 /6	31-35	THIRD CARD IMAGE
Base/Runways	"25th ARC #"	CT 17 /6	36-40	THIRD CARD IMAGE
Base/Runways	"26th ARC #"	CT 17 /6	41-45	THIRD CARD IMAGE
Base/Runways	"27th ARC #"	CT 17 /6	46-50	THIRD CARD IMAGE
Base/Runways	"28th ARC #"	CT 17 /6	51-55	THIRD CARD IMAGE
Base/Runways	"29th ARC #"	CT 17 /6	56-60	THIRD CARD IMAGE
Base/Runways	"30th ARC #"	CT 17 /6	61-65	THIRD CARD IMAGE
Base/Runways	"31th ARC #"	CT 17 /6	16-20	THIRD CARD IMAGE
Base/Runways	"32th ARC #"	CT 17 /6	21-25	THIRD CARD IMAGE
Base/Runways	"33th ARC #"	CT 17 /6	26-30	THIRD CARD IMAGE
Base/Runways	"34th ARC #"	CT 17 /6	31-35	THIRD CARD IMAGE
Base/Runways	"35th ARC #"	CT 17 /6	35-40	THIRD CARD IMAGE
Base/Runways	"36th ARC #"	CT 17 /6	41-45	THIRD CARD IMAGE
Base/Runways	"37th ARC #"	CT 17 /6	46-50	THIRD CARD IMAGE
Base/General	"RRNODE"	CT 17 /7	11-15	
Base/General	"Extended MCL"	CT 17 /7	16-20	
Base/General	"Extended MCW"	CT 17 /7	21-25	
Base/General	"Max Runways"	CT 17 /7	26-30	
Base/General	"MCL"	CT 17 /7	31-35	
Base/General	"MCW"	CT 17 /7	36-40	

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Base/Ramps	"Nearest Node"	CT 17 /8	21-25	CC 15 = 2
Base/Ramps	"Relative Cap."	CT 17 /8	21-25	CC 15 = 1
Base/General	"PA Task Delay"	CT 17 /9	16-25	
Base/General	"Survey /EOD Time"	CT 17 /9	26-30	
Base/General	"Begin Day Shift"	CT 18 /1	VAR	SHOP DEPENDENT
Scenario/General	"Unsh Mnt Prb Mod"	CT 18 /2		
Base/AC_Basing	"Aircraft Name"	CT 20	6-9	
Base/AC_Basing	"Quantity"	CT 20	11-15	
Base/AC_Basing	"Arrival Time"	CT 20 /66	6-10	NOTE UNITS
Scenario/AC_Spares	"Aircraft Name"	CT 20 /77	10	
Scenario/AC_Spares	"Quantity"	CT 20 /77	11-15	
Scenario/AC_Spares	"Reorder Hours"	CT 20 /77	16-20	UNITS NDE
Base/Personnel	"Initial Number"	CT 21	VAR	NDE
Base/Personnel	"Target Number"	CT 21	VAR	NDE
Base/Personnel	"% Day Shift"	CT 21	VAR	NDE
Base/Personnel	"Min Crew Size"	CT 21	VAR	NDE
Base/Personnel	"Personnel Name"	CT 21	VAR	NDE
Base/AGE	"AGE Name"	CT 22	6-10	
Base/AGE	"Initial Number"	CT 22	12-13	
Base/Parts	"Part Name"	CT 23	6-10	
Base/Parts	"Initial Number"	CT 23	12-13	
Base/Parts	"Min Inventory"	CT 23	21-25	
Resource/Resources	"Base NRTS Rate"	CT 23 /2XX	VAR	26-30 THRU 76-80
Base/Munitions	"Munition Name"	CT 24	VAR	6-10 THRU 66-70
Base/Munitions	"Initial Number"	CT 24	VAR	11-15 THRU 71-75
Base/TRAP	"TRAP Name"	CT 25	VAR	6-10 THRU 46-50
Base/TRAP	"Initial Number"	CT 25	VAR	11-15 THRU 51-55
Base/General	"Init POL Stocks"	CT 27	11-15	
Base/AC_Networks	"Task Field 1"	CT 29	16-20	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 2"	CT 29	21-25	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 3"	CT 29	26-30	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 4"	CT 29	31-35	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 5"	CT 29	36-40	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 6"	CT 29	41-45	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 7"	CT 29	46-50	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 8"	CT 29	51-55	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 9"	CT 29	56-60	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 10"	CT 29	61-65	FIRST CARD IMAGE
Base/AC_Networks	"Task Field 11"	CT 29	16-20	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 12"	CT 29	21-25	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 13"	CT 29	26-30	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 14"	CT 29	31-35	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 15"	CT 29	36-40	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 16"	CT 29	41-45	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 17"	CT 29	46-50	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 18"	CT 29	51-55	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 19"	CT 29	56-60	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 20"	CT 29	61-65	SECOND CARD IMAGE
Base/AC_Networks	"Task Field 21"	CT 29	16-20	THIRD CARD IMAGE

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Base/AC_Networks	"Task Field 22"	CT 29	21-25	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 23"	CT 29	26-30	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 24"	CT 29	31-35	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 25"	CT 29	36-40	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 26"	CT 29	41-45	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 27"	CT 29	46-50	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 28"	CT 29	51-55	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 29"	CT 29	56-60	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 30"	CT 29	61-65	THIRD CARD IMAGE
Base/AC_Networks	"Task Field 31"	CT 29	16-20	FOURTH CARD IMAGE
Base/AC_Networks	"Task Field 32"	CT 29	21-25	FOURTH CARD IMAGE
Base/AC_Networks	"Task Field 33"	CT 29	26-30	FOURTH CARD IMAGE
Base/AC_Networks	"Task Field 34"	CT 29	31-35	FOURTH CARD IMAGE
Base/AC_Networks	"Task Field 35"	CT 29	36-40	FOURTH CARD IMAGE
Base/AC_Networks	"Task Field 36"	CT 29	41-45	FOURTH CARD IMAGE
Base/AC_Networks	"Task Field 37"	CT 29	46-50	FOURTH CARD IMAGE
Base/General	"POL Reord Amt"	CT 31	VAR	TYPE 0
Base/Parts	"Reorder Quantity"	CT 31	VAR	26-30 THRU 71-75
Base/Munitions	"Reorder Quantity	CT 31	VAR	26-30 THRU 71-75
Base/TRAP	"Reorder Quantity"	CT 31	VAR	26-30 THRU 71-75
Res_Order/Res_Order	"Resource Name"	CT 31	VAR	21-25 THRU 66-70
Res_Order/Res_Order	"Resource Type*"	CT 31	VAR	31-35 THRU 76-80
Res_Order/Res_Order	"Day"	CT 31	6-10	
Res_Order/Rtype_Order	"POL Arriv Time"	CT 31	6-15	NDE
Res_Order/Rtype_Order	"MUN Arriv Time"	CT 31	6-15	NDE
Res_Order/Res_Order	"Time-to-Arrive"	CT 31	11-15	
Base/AGE	"Reorder Quantity	CT 31	VAR	26-30 THRU 71-75
Res_Order/Rtype_Order	"PART Dist Param"	CT 33		
Res_Order/Rtype_Order	"PART Dist Type*"	CT 33		
Scenario/General	"Auto Res Rspy?*"	CT 33	NDE	
Res_Order/Rtype_Order	"AGE Dist Type*"	CT 33	VAR	16-20 TO 76-80 NDE
Res_Order/Rtype_Order	"AGE Dist Param"	CT 33	VAR	16-20 TO 76-80 NDE
Res_Order/Rtype_Order	"TRAP Arriv Time"	CT 33	VAR	11-15 TO 71-75 NDE
Res_Order/Rtype_Order	"PART Arriv Time"	CT 33	VAR	11-15 TO 71-75 NDE
Res_Order/Rtype_Order	"AGE Arriv Time"	CT 33	VAR	11-15 TO 71-75 NDE
Res_Order/Rtype_Order	"TRAP Dist Param"	CT 33	VAR	16-20 TO 76-80 NDE
Res_Order/Rtype_Order	"TRAP Dist Type*"	CT 33	VAR	16-20 TO 76-80 NDE
Base/General	"Crater Rep Time"	CT 38	11-13	NDE
Scenario/AC_Spares	"Init Flight Hrs"	CT 41	SUM	NDE
		CT 42	SUM	NDE
Scenario/AC_Spares	"Initial Status*"	CT 41	SUM	NDE
		CT 42	SUM	NDE
Base/AC_Basing	"Initial Status*"	CT 41		NDE
Base/AC_Basing	"Mission Config"	CT 41		NDE
Base/Cross_Train	"Requested Type"	CT 45/2	6-10	
Base/Cross_Train	"CT Type # 1"	CT 45/2	11-15	
Base/Cross_Train	"CT Type # 2"	CT 45/2	16-20	
Base/Cross_Train	"CT Type # 3"	CT 45/2	21-25	
Base/Cross_Train	"CT Type # 4"	CT 45/2	26-30	

SORGEN DATABASE/RELATION	SORGEN DATA LABEL	TSAR CARD TYPE	CARD COLUMN	COMMENT
Base/Cross_Train	"CT Type # 5"	CT 45/2	31-35	
Scenario/Missions	"Lst Takeoff Time"	CT 50		MDE
Scenario/Missions	"Type*"	CT 50		MDE
Scenario/Missions	"1st Takeoff Time"	CT 50		MDE
Mission/Mission	"Aircraft Type"	CT 50	11-15	
Scenario/Missions	"Mission Name"	CT 50	16-20	MDE
Mission/Mission	"Priority"	CT 50	21-25	
Mission/Mission	"Desired # AC"	CT 50	31-35	
Mission/Mission	"Minimum # AC"	CT 50	36-40	
Scenario/Missions	"Hours Notice"	CT 50	41-45	

Appendix F: TSAR Database Changes Required for Equilibration

	Change to TSAR Database	Reason for Change
1.	Air traffic control (ATC) disabled	SORGEN does not simulate ATC
2.	Cannibalization of aircraft parts disabled	SORGEN does not simulate cannibalization
3.	Emergency base disabled	Single base simulation
4.	Task interrupt disabled	SORGEN does not simulate task interrupt
5.	Fuel task mutually exclusive	SORGEN fueling task fixed as mutually exclusive
6.	Munition assembly disabled	SORGEN does not simulate munition assembly
7.	Alternative mission configurations disabled	SORGEN simulates a only a single configuration per mission type
8.	Standard combat loads reduced to one per mission type	SORGEN simulates only one combat load per mission type
9.	Basic munitions load disabled	SORGEN does not simulate basic munitions load, only mission specific munitions
10.	Begin day shift set to 0600 for all shops	SORGEN allows only a single begin day shift entry for use by all shops
11.	Delete all munition components replace with equivalent built up munitions	SORGEN does not simulate munition build up tasks
12.	Delete all CT43 and CT44 weather and chemical warfare data entries in database	SORGEN does not simulate weather or chemical warfare
13.	Delete all CT5 columns 58-60 alternative task identification number	SORGEN does not simulate alternative procedures when personnel or support equipment are unavailable
14.	Delete all CT6 alternative task data entries in database	SORGEN does not simulate alternative procedures
15.	Aggregate CT8/2 and CT8/3 multiple step component repair procedures	SORGEN does not simulate multiple step off equipment repair actions on components
16.	Delete CT23/66 cost data entries	Cost data not being used in simulation
17.	Delete CT23/7X spares computation data entries	Spares computation feature ,OUTFIT, of TSAR not used, no equivalent in SORGEN

	Change to TSAR Database	Reason for Change
18.	Delete CT28 part salvage data entries	SORGEN does not simulate salvage of components from damaged aircraft
19.	Delete CT35 part salvage data entries	SORGEN does not simulate salvage of components from damaged aircraft
20.	Delete CT16/88 and CT16/99 missions flown attrition adjustment data entries	Attrition modification based on missions flown not used, research experiment uses daily attrition rate
21.	Aggregate personnel types for use in SORGEN. Leave in combat oriented maintenance organization (COMO) configuration for TSAR	SORGEN does not simulate COMO maintenance structure
22.	Combine Type 53 and Type 80 AGE for use in SORGEN. Leave as is for TSAR	SORGEN uses fixed refueling task, aggregation needed for equal refueling resources in both models
23.	Calculate expected value of increased component repair time by AIS type	Add to parts using AIS AGE types to account for maintenance time on stations per use. SORGEN does not simulate support equipment repair as a function of aircraft part repair
24.	Calculate expected value of probability of awaiting part for AIS repair	Use product of expected value times part delay for repair time for AIS repair in SORGEN. SORGEN does not simulate awaiting parts without LOGSIM
25.	Delete all base two personnel, spares, and support equipment entries	SORGEN simulates multiple bases however, research experiment simulates a single base
26.	Aggregate CT13 and CT14 mission preparation times for use in SORGEN	SORGEN takes a single mission preparation task which includes TRAP and mission specific munitions
27.	Delete TRAP configuration	SORGEN does not track aircraft TRAP configuration. TRAP is removed after each mission.
28.	Calculate fuel requirements for fuel treatments using 12 unit quantity in TSAR (setting used in database)	Calculate fuel requirements for fuel treatments using 10 unit quantity in SORGEN, value fixed in fueling task
29.	Set munition retention to 0% on CT16	SORGEN assumes all munitions expended on each sortie
30.	Delete CT8/3 column 36-40 data entries to disable SRU repair	SORGEN does not simulate SRU repair actions
31.	Disable simulation of air crews	SORGEN does not simulate air crews

Appendix G: SAS Output for Variability Runs

Data for Test of Normality

OBS	TSAR H/30/30	TSAR L/30/30	SORGEN H/30/30	SORGEN L/30/30
1	3322	3077	2784	1877
2	3499	2666	2975	1686
3	3967	3053	3246	2018
4	2290	2906	2822	1661
5	5099	3121	2723	2030
6	2458	3064	2898	1930
7	4967	2892	3119	1842
8	3329	2514	3336	1830
9	3769	3092	2759	1883
10	3363	2863	2716	1920
11	1894	2720	3068	1953
12	4813	2924	2861	1743
13	3098	2943	2877	1938
14	4938	2365	2972	1841
15	2861	3191	2951	1653
16	4572	2876	2918	1861
17	3857	3223	3011	1745
18	4585	3230	3001	2048
19	4642	3202	3068	1782
20	3848	2960	3023	1578
21	3194	2970	2894	1891
22	5146	3055	2943	1936
23	3784	3008	2916	1735
24	5191	3056	3214	1880
25	3426	2894	3109	2064
26	2811	3002	3227	1684
27	3776	2869	3229	1495
28	3885	2782	2824	1929
29	2087	2954	2609	1893
30	3873	2990	3135	1875

Normality Test TSAR High Treatment-Thirty Trials

Univariate Procedure

Variable=T_HI30

Moments

N	30	Sum Wgts	30
Mean	3744.8	Sum	112344
Std Dev	932.8224	Variance	870157.7
Skewness	-0.15458	Kurtosis	-0.70348
USS	4.4594E8	CSS	25234573
CV	24.90981	Std Mean	170.3093
T:Mean=0	21.98823	Prob>|T|	0.0001
Num ^= 0	30	Num > 0	30
M(Sign)	15	Prob>|M|	0.0001
Sgn Rank	232.5	Prob>|S|	0.0001
W:Normal	0.953395	Prob<W	0.2416

Quantiles(Def=5)

100% Max	5191	99%	5191
75% Q3	4585	95%	5146
50% Med	3780	90%	5033
25% Q1	3194	10%	2374
0% Min	1894	5%	2087
		1%	1894
Range	3297		
Q3-Q1	1391		
Mode	1894		

Extremes

Lowest	Obs	Highest	Obs
1894(	11)	4938(	14)
2087(	29)	4967(	7)
2290(	4)	5099(	5)
2458(	6)	5146(	22)
2811(	26)	5191(	24)

```
Stem Leaf                     #             Boxplot
   5 0112                     4                |
   4 66689                    5             +-----+
   4 0                        1             |     |
   3 58888999                 8             *--+--*
   3 123344                   6             +-----+
   2 589                      3                |
   2 13                       2                |
   1 9                        1                |
     ----+----+----+----+
 Multiply Stem.Leaf by 10**+3
```

Normality Test, TSAR High Treatment-Thirty Trials

Univariate Procedure

Variable=T_HI30

```
                    Normal Probability Plot
  5250+                                            *+*++ *
      |                                    ******+*+
      |                                ++++++
      |                          +*******
      |                   *******
      |            +++**+
      |      ++*+* *
  1750+ +++++*+
       +----+----+----+----+----+----+----+----+----+----+
           -2        -1         0        +1        +2
```

Normality Test, TSAR Low Treatment-Thirty Trials

Univariate Procedure

Variable=T_LO30

Moments

N	30	Sum Wgts	30
Mean	2948.733	Sum	88462
Std Dev	196.2509	Variance	38514.41
Skewness	-1.13978	Kurtosis	1.910144
USS	2.6197E8	CSS	1116918
CV	6.65543	Std Mean	35.83035
T:Mean=0	82.2971	Prob>|T|	0.0001
Num ^= 0	30	Num > 0	30
M(Sign)	15	Prob>|M|	0.0001
Sgn Rank	232.5	Prob>|S|	0.0001
W:Normal	0.918815	Prob<W	0.0281

Quantiles(Def=5)

100% Max	3230	99%	3230
75% Q3	3064	95%	3223
50% Med	2965	90%	3196.5
25% Q1	2876	10%	2693
0% Min	2365	5%	2514
		1%	2365
Range	865		
Q3-Q1	188		
Mode	2365		

Extremes

Lowest	Obs	Highest	Obs
2365(	14)	3121(	5)
2514(	8)	3191(	15)
2666(	2)	3202(	19)
2720(	11)	3223(	17)
2782(	28)	3230(	18)

```
Stem Leaf                     #             Boxplot
  32 023                      3                |
  31 29                       2                |
  30 01566689                 8             +-----+
  29 1245679                  7             *--+--*
  28 67899                    5             +-----+
  27 28                       2                |
  26 7                        1                |
  25 1                        1                0
  24
  23 6                        1                0
     ----+----+----+----+
 Multiply Stem.Leaf by 10**+2
```

Normality Test, TSAR Low Treatment-Thirty Trials

Univariate Procedure

Variable=T_LO30

```
                   Normal Probability Plot
 3250+                                          ++*+*+   *
     |                                     ++*+*
     |                               ********
 2950+                         ******+
     |                    *****++
     |                *+*++
 2650+           ++++*+
     |      +++++ *
     |+++
 2350+     *
      +----+----+----+----+----+----+----+----+----+----+
          -2        -1         0        +1        +2
```

Normality Test, SORGEN High Treatment-Thirty Trials

Univariate Procedure

Variable=S_HI30

Moments

N	30	Sum Wgts	30
Mean	2974.267	Sum	89228
Std Dev	177.3434	Variance	31450.69
Skewness	0.105624	Kurtosis	-0.47005
USS	2.663E8	CSS	912069.9
CV	5.962593	Std Mean	32.37833
T:Mean=0	91.8598	Prob>|T|	0.0001
Num ^= 0	30	Num > 0	30
M(Sign)	15	Prob>|M|	0.0001
Sgn Rank	232.5	Prob>|S|	0.0001
W:Normal	0.984777	Prob<W	0.9391

Quantiles(Def=5)

100% Max	3336	99%	3336
75% Q3	3109	95%	3246
50% Med	2961.5	90%	3228
25% Q1	2861	10%	2741
0% Min	2609	5%	2716
		1%	2609
Range	727		
Q3-Q1	248		
Mode	3068		

Extremes

Lowest	Obs	Highest	Obs
2609(	29)	3214(	24)
2716(	10)	3227(	26)
2723(	5)	3229(	27)
2759(	9)	3246(	3)
2784(	1)	3336(	8)

```
Stem Leaf                    #           Boxplot
  33 4                       1              |
  32 1335                    4              |
  31 124                     3           +-----+
  30 01277                   5           |     |
  29 0224578                 7           *--+--*
  28 22689                   5           +-----+
  27 2268                    4              |
  26 1                       1              |
     ----+----+----+----+
 Multiply Stem.Leaf by 10**+2
```

Normality Test, SORGEN High Treatment-Thirty Trials

Univariate Procedure

Variable=S_HI30

```
                       Normal Probability Plot
  3350+                                              +*++++
      |                                      * *+*+*++
      |                                   *****++
      |                              *****++
      |                        +*****
      |                   ******
      |             *+*+*+*
  2650+     +*++++
       +----+----+----+----+----+----+----+----+----+----+
           -2        -1         0        +1        +2
```

Normality Test, SORGEN Low Treatment-Thirty Trials

Univariate Procedure

Variable=S_LO30

Moments

N	30	Sum Wgts	30
Mean	1840.033	Sum	55201
Std Dev	139.307	Variance	19406.45
Skewness	-0.59409	Kurtosis	0.021678
USS	1.0213E8	CSS	562787
CV	7.570897	Std Mean	25.43387
T:Mean=0	72.3458	Prob>|T|	0.0001
Num ^= 0	30	Num > 0	30
M(Sign)	15	Prob>|M|	0.0001
Sgn Rank	232.5	Prob>|S|	0.0001
W:Normal	0.95608	Prob<W	0.2824

Quantiles(Def=5)

100% Max	2064	99%	2064
75% Q3	1930	95%	2048
50% Med	1876	90%	2024
25% Q1	1743	10%	1657
0% Min	1495	5%	1578
		1%	1495
Range	569		
Q3-Q1	187		
Mode	1495		

Extremes

Lowest	Obs	Highest	Obs
1495(	27)	1953(	11)
1578(	20)	2018(	3)
1653(	15)	2030(	5)
1661(	4)	2048(	18)
1684(	26)	2064(	25)

```
Stem Leaf                    #             Boxplot
  20 56                      2                |
  20 23                      2                |
  19 5                       1                |
  19 23344                   5             +-----+
  18 6888899                 7             *-----*
  18 344                     3             |  +  |
  17 8                       1             |     |
  17 444                     3             +-----+
  16 5689                    4                |
  16                                          |
  15 8                       1                |
  15 0                       1                |
  14
     ----+----+----+----+
 Multiply Stem.Leaf by 10**+2
```

Normality Test, SORGEN Low Treatment-Thirty Trials

Univariate Procedure

Variable=S_LO30

```
                          Normal Probability Plot
   2075+                                                  +++ *
       |                                              *+*+*
       |                                          ++*+
       |                                       *****
       |                                 ******
       |                              ****++
   1775+                            +*++
       |                        ++***
       |                    *+*+**
       |                 ++++
       |             +++ *
       |         ++++
   1475++    *
        +----+----+----+----+----+----+----+----+----+----+
            -2        -1         0        +1        +2
```

Appendix H: SAS Output for Experimental Runs

Data for Paired Difference Test - Experimental Runs

(Sorties per 30 days)

Treatment	SORGEN Results	TSAR Results	Paired Differences
1	1842.3	2977.9	1135.6
2	2758.1	4501.9	1743.8
3	1863.2	3373.2	1510.0
4	2825.1	3984.1	1159.0
5	2471.2	3540.2	1069.0
6	2103.3	3215.8	1112.5
7	2326.9	3264.8	937.9
8	2146.8	3958.7	1811.9
9	1911.3	3417.4	1506.1
10	2824.5	3679.4	854.9
11	1870.2	2745.5	875.3
12	2744.5	3461.2	716.7
13	2403.4	2759.3	355.9
14	2197.9	3968.6	1770.7
15	2471.4	3724.7	1253.3
16	2106.6	3187.4	1080.8
17	2163.1	3745.4	1582.3
18	2410.8	3438.4	1027.6
19	2108.8	3062.1	953.3
20	2467.4	2994.1	526.7
21	2776.9	3415.8	638.9
22	1860.3	3276.8	1416.5
23	2812.0	3984.1	1172.1
24	1797.7	2956.7	1159.0
25	2089.4	3023.5	934.1
26	2543.4	3821.8	1278.4
27	2142.5	3729.7	1587.2
28	2515.6	3080.7	565.1
29	2966.4	3739.1	772.7
30	1873.1	2741.0	867.9
31	2691.2	3679.4	988.2
32	1859.7	3375.2	1515.5
33	2458.6	2649.1	190.5
34	2166.4	3668.5	1502.1
35	2495.9	3723.4	1227.5
36	2135.1	3157.8	1022.7
37	1862.4	3541.1	1678.7
38	2772.9	3508.7	735.8
39	1887.0	2951.7	1064.7
40	2765.1	3464.0	698.9
41	2532.6	3450.3	917.7
42	2120.3	3023.1	902.8
43	2427.1	3210.4	783.3

44	2186.6	4163.5	1976.9
45	1887.0	2954.5	1067.5
46	2753.1	4030.5	1277.4
47	1866.6	3245.6	1379.0
48	2780.8	2817.1	36.3
49	2984.6	3566.9	582.3
50	1829.3	2728.8	899.5
51	2704.8	3642.4	937.6
52	1863.5	3365.8	1502.3
53	2077.2	3247.8	1170.6
54	2473.1	4039.9	1566.8
55	2146.4	3541.7	1395.3
56	2468.5	3212.5	744.0
57	2763.9	3352.3	588.4
58	1809.5	3226.6	1417.1
59	2896.2	3798.1	901.9
60	1859.2	2773.4	914.2
61	2145.6	3771.6	1626.0
62	2478.9	3434.9	956.0
63	2088.9	2937.9	849.0
64	2467.4	2815.8	348.4

Paired Difference Test

Univariate Procedure

Variable=TRTDIFF

Moments

N	64	Sum Wgts	64
Mean	1074.064	Sum	68740.1
Std Dev	407.9631	Variance	166433.9
Skewness	-0.08235	Kurtosis	-0.15004
USS	84316608	CSS	10485337
CV	37.98313	Std Mean	50.99539
T:Mean=0	21.06198	Prob>|T|	0.0001
Num ^= 0	64	Num > 0	64
M(Sign)	32	Prob>|M|	0.0001
Sgn Rank	1040	Prob>|S|	0.0001

Quantiles(Def=5)

100% Max	1976.9	99%	1976.9
75% Q3	1405.9	95%	1743.8
50% Med	1046.15	90%	1587.2
25% Q1	851.95	10%	582.3
0% Min	36.3	5%	355.9
		1%	36.3
Range	1940.6		
Q3-Q1	553.95		
Mode	1159		

Extremes

Lowest	Obs	Highest	Obs
36.3(	48)	1678.7(	37)
190.5(	33)	1743.8(	2)
348.4(	64)	1770.7(	14)
355.9(	13)	1811.9(	8)
526.7(	20)	1976.9(	44)

Normality Test, Treatment Differences

Univariate Procedure

Variable=TRTDIFF

Moments

N	64	Sum Wgts	64
Mean	1074.064	Sum	68740.1
Std Dev	407.9631	Variance	166433.9
Skewness	-0.08235	Kurtosis	-0.15004
USS	84316608	CSS	10485337
CV	37.98313	Std Mean	50.99539
T:Mean=0	21.06198	Prob>|T|	0.0001
Num ^= 0	64	Num > 0	64
M(Sign)	32	Prob>|M|	0.0001
Sgn Rank	1040	Prob>|S|	0.0001
W:Normal	0.98262	Prob<W	0.7669

Quantiles(Def=5)

100% Max	1976.9	99%	1976.9
75% Q3	1405.9	95%	1743.8
50% Med	1046.15	90%	1587.2
25% Q1	851.95	10%	582.3
0% Min	36.3	5%	355.9
		1%	36.3
Range	1940.6		
Q3-Q1	553.95		
Mode	1159		

Extremes

Lowest	Obs	Highest	Obs
36.3(	48)	1678.7(	37)
190.5(	33)	1743.8(	2)
348.4(	64)	1770.7(	14)
355.9(	13)	1811.9(	8)
526.7(	20)	1976.9(	44)

Normality Test, Treatment Differences
Univariate Procedure

Variable=TRTDIFF

```
Stem Leaf                     #             Boxplot
  19 8                        1                |
  18 1                        1                |
  17 47                       2                |
  16 38                       2                |
  15 00112789                 8                |
  14 022                      3             +-----+
  13 8                        1             |     |
  12 3588                     4             |     |
  11 146677                   6             |     |
  10 236778                   6             *--+--*
   9 00012344569             11             |     |
   8 5578                     4             +-----+
   7 024478                   6                |
   6 4                        1                |
   5 3789                     4                |
   4                                          |
   3 56                       2                |
   2                                          |
   1 9                        1                |
   0 4                        1                |
     ----+----+----+----+
Multiply Stem.Leaf by 10**+2
```

```
                   Normal Probability Plot
1950+                                              ++*
    |                                           ++*
    |                                         **
    |                                      +**
    |                                 ******
    |                                *+++
    |                               *+
    |                             ***
    |                          +***
    |                        ***
    |                    *****
    |                  ***
    |               ****
    |              **
    |          ****
    |         ++
    |       +**
    |    +++
    |  ++ *
  50+++*
     +----+----+----+----+----+----+----+----+----+----+
         -2        -1         0        +1        +2
```

Bibliography

1. Balci, Osman. "How to Assess the Acceptability and Credibility of Simulation Results," *Proceedings of the 1989 Winter Simulation Conference*. 62-71. Englewood Cliffs NJ: IEEE Press, 1989.

2. Ball Systems Engineering Division. *All Mobile Tactical Air Force (AMTAF) System User's Manual.* R-024-89. San Diego CA: BSED, March 1989.

3. Ball Systems Engineering Division. *All Mobile Tactical Air Force (AMTAF) System User's Manual Appendices.* R-024-89. San Diego CA: BSED, March 1989.

4. Ball Systems Engineering Division. *All Mobile Tactical Air Force (AMTAF) System Programmer/Analyst's Manual.* R-025-89. San Diego CA: BSED, March 1989.

5. Ball Systems Engineering Division. *All Mobile Tactical Air Force (AMTAF) System Programmer/Analyst's Manual Appendices.* R-025-89. San Diego CA: BSED, March 1989.

6. Ball Systems Engineering Division. *VICS User's Guide.* R-005-90. Boulder CO: BSED, April 1990.

7. Ball Systems Engineering Division. *Modeling of Reliability and Maintainability (R&M) Goals for Tactical Aircraft.* Final Report, R-033-88. San Diego CA: BSED, May 1988.

8. Carson, John. "Verification and Validation: A Consultant's Perspective," *Proceedings of the 1989 Winter Simulation Conference*. 552-558. Englewood Cliffs NJ: IEEE Press, 1989.

9. Clark, Capt Gregg A. *The Theater Simulation of Airbase Resources and Logistics Composite Models: A Comparison.* MS Thesis, AFIT/GLM/LSM/87S-15. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1987 (AD-A188045).

10. Cook, Thomas M. and Robert A. Russell. *Introduction to Management Science.* Englewood Cliffs NJ: Prentice Hall, 1989.

11. Department of the Air Force. *Compendium of Authenticated Systems and Logistics Terms, Definitions and Acronyms.* AU-AFIT-LS-3-81. School of Systems & Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, April 1981.

12. Diener, David A. *Forecasting Air Base Operability in a Hostile Environment: Estimating Metamodels From Large Scale Simulations.* Unpublished PhD dissertation. Purdue University, West Lafayette IN, December 1989.

13. Diener, Lt Col David A. Professor, Department of Logistics Management. Personal Interviews. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1991 - July 1992.

14. Emerson, Donald E. *An Introduction to the TSAR Simulation Program - Model Features and Logic.* R-2584-AF. Santa Monica CA: Rand Corp., February 1982.

15. Emerson, Donald E. *TSAR User's Manual-A Program for Assessing the Effects of Conventional and Chemical Attacks on Sortie Generation: Vol. II, Data Input, Program Operation and Redimensioning, and Sample Problem.* N-3012-AF. Santa Monica CA, RAND Corp., September 1990.

16. Emerson, Donald E. *TSAR User's Manual-A Program for Assessing the Effects of Conventional and Chemical Attacks on Sortie Generation: Vol. III, Variable and Array Definitions, and Other Program Aids.* N-3013-AF. Santa Monica CA, RAND Corp., September 1990.

17. Emerson, Donald E. and Louis H. Wegner. *TSARINA-A Computer Model for Assessing Conventional and Chemical Attacks on Airbases.* N-3010-AF. Santa Monica CA, RAND Corp., September 1990.

18. Emerson, Donald E. and Louis H. Wegner. *TSAR User's Manual - A Program for Assessing the Effects of Conventional and Chemical Attacks on Sortie Generation: Vol. I, Program Features, Logic and Interactions.* N-3011-AF. Santa Monica CA, RAND Corp., September 1990.

19. Fossett, Christine A., Dale Harrison, Harry Weintrob and Saul I. Gass. "An Assessment Procedure For Simulation Models: A Case Study," *Operations Research, 39:* 710-723 (September-October 1991).

20. Gass, Saul I. "Managing the Modeling Process: A Personal Reflection," *European Journal of Operations Research, 31*: 1-8 (1987).

21. General Accounting Office. *DOD SIMULATIONS Improved Assessment Procedures Would Increase the Credibility of Results.* GAO/PEMD-88-3. Washington DC: Government Printing Office, 1987.

22. General Accounting Office. *Guidelines for Model Evaluations.* PAD-79-17. Washington DC: Government Printing Office, 1979.

23. Law, Averill M. and W. David Kelton. *Simulation Modeling and Analysis*. New York: McGraw-Hill, 1982.

24. Law, Averill M. and W. David Kelton. *Simulation Modeling and Analysis* (Second Edition). New York: McGraw-Hill, 1991.

25. Leibholz, Stephen W. "Twenty Questions," *Military Modeling*, edited by Wayne P. Hughes, Jr. Alexandria VA: Military Operations Research Society, 1984.

26. Leonhardt, Capt David P. *A Comparison of the All Mobile Tactical Air Force and the Logistics Composite Simulation Models*. MS Thesis, AFIT/GLM/LSM/91S-42. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1991 (AD-A246761).

27. McClave, James T. and P. George Benson. *Statistics for Business and Economics*. New York: MacMillan Publishing Company, 1991.

28. McLean, Robert A. and Virgil L. Anderson. *Applied Factorial and Fractional Designs*. New York: Marcel Dekker Incorporated, 1984.

29. Noble, Capt David R. *Comparison of the TSAR Model to the LCOM Model*. MS thesis, AFIT/GLM/LSM/86S-54. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1986 (AD-A174330).

30. Shannon, Robert E. *Systems Simulation, The Art and Science*. Englewood Cliffs NJ: Prentice-Hall, Inc., 1975.

31. Van Horn, Richard L. "Validation of Simulation Results," *Management Science, 17*: 247-258 (January 1971).

32. Williams, Marion L. and James Sikora. "SIMVAL Minisymposium - A Report," *Phalanx, 24*: 1-6 (June 1991).

Vita

Captain Heston R. Hicks was born 10 October 1961 at Fort Bragg, North Carolina. He graduated from Graceville High School in Graceville, Florida in 1979 before receiving his Bachelor of Science degree in Biology from Florida State University in 1983. In 1985, he graduated from the Air Force Officer Training School and entered active duty. His first logistics assignment was at the 91st Strategic Missile Wing, Minot AFB, North Dakota where he served as the War Reserve Materiel Officer, Installation Mobility Officer, Assistant Chief, Logistics Plans and Programs Division, and Chief, Logistics Plans and Programs Division. In 1988, Captain Hicks was assigned to Headquarters, Air Force Systems Command, Andrews AFB, Maryland. While assigned there he served as Acquisition Logistics Program Integration Manager to the Deputy Chief of Staff for Product Assurance and Acquisition Logistics, and as Engineering and Technical Project Integration Manager to the Deputy Chief of Staff for Engineering and Technical Management. He entered the School of Systems and Logistics, Air Force Institute of Technology in May 1991.

Permanent Address: 670 Satellite Road
Graceville, Florida 32440

Vita

Mr. Lawrence L. Long was born 4 April 1948 in Gunnison, Colorado. He graduated from Telluride High School in 1966 and attended Western State College of Colorado for one year, prior to entering active duty in the United States Air Force. While on active duty Mr. Long served as an electronic technician. He was honorably discharged on 12 June 1974 after serving tours at Keesler AFB, MS, Otis AFB, MA, Blytheville AFB, AR, Ubon and Korat, Thailand, and Minot AFB, ND. Following active duty he was employed as an Air Force Engineering and Technical Service representative to the 5th Bombardment Wing at Minot AFB, ND. In August 1983 Mr. Long accepted an assignment to SAC headquarters in Omaha, NE where he served as a electronic warfare staff officer for the Deputy Chief of Staff for Logistics. In 1988 he received his Bachelor of Science Degree from the Offutt AFB Extension of Southern Illinois University - Carbondale. In March 1989, Mr. Long accepted a logistics management position with the Air Force Electronic Combat Office, Wright-Patterson AFB, OH where he served until entering the School of Systems and Logistics, Air Force Institute of Technology in May 1991.

Permanent address: 1431 Penbrooke Trail
Centerville, OH 45459-3337

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1992	Master's Thesis

4. TITLE AND SUBTITLE
A METHODOLOGY FOR MODEL COMPARISON USING THE THEATER SIMULATION OF AIRBASE RESOURCES AND ALL MOBILE TACTICAL AIR FORCE MODELS

5. FUNDING NUMBERS

6. AUTHOR(S)
Heston R. Hicks, Captain, USAF
Lawrence L. Long, Civilian, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology, WPAFB OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GLM/LSM/92S-24

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Eric Werkowitz
HQ AFMC/STXP
Wright-Patterson AFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*) This research compares the Theater Simulation of Airbase Resources (TSAR) model to the sortie generation (SORGEN) module of the All Mobile Tactical Air Force (AMTAF) model, qualitatively and quantitatively, while concurrently developing and proving a model comparison methodology. The qualitative analysis compares the models' background and documentation, features and databases, and useability. The quantitative analysis statistically compares the models' estimates of sorties generated. For the quantitative study, eight variables are chosen and assigned high and low values for use in a 2^8 1/4 fractional factorial experimental design. Equivalent input databases are developed from a TSAR F-15C database and pilot trials are run to test the factor levels and assess variability. Finally, 64 experimental trials are run and paired differences of the results are tested to determine the statistical equivalence of the models. Results reveal notable differences in the models, both qualitative and quantitative. Further research is needed to analyze the quantitative differences. The qualitative differences are believed due primarily to differences in the models' designed fidelity. The methodology developed provides a functional framework for model comparison and is improved for use in future research.

14. SUBJECT TERMS
Model Comparison, Credibility, Acceptance, Accreditation, Simulation, Theater Simulation of Airbase Resources, All Mobile Tactical Air Force

15. NUMBER OF PAGES
183

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

AFIT Control Number AFIT/GLM/LSM/92S-24

AFIT RESEARCH ASSESSMENT

The purpose of this questionnaire is to determine the potential for current and future applications of AFIT thesis research. Please return completed questionnaires to: AFIT/LSC, Wright-Patterson AFB OH 45433-9905.

1. Did this research contribute to a current research project?

   a. Yes b. No

2. Do you believe this research topic is significant enough that it would have been researched (or contracted) by your organization or another agency if AFIT had not researched it?

   a. Yes b. No

3. The benefits of AFIT research can often be expressed by the equivalent value that your agency received by virtue of AFIT performing the research. Please estimate what this research would have cost in terms of manpower and/or dollars if it had been accomplished under contract or if it had been done in-house.

   Man Years ______________ $ ______________

4. Often it is not possible to attach equivalent dollar values to research, although the results of the research may, in fact, be important. Whether or not you were able to establish an equivalent value for this research (3, above) what is your estimate of its significance?

   a. Highly Significant b. Significant c. Slightly Significant d. Of No Significance

5. Comments

______________________ ______________________

Name and Grade Organization

______________________ ______________________

Position or Title Address